AD_____

Award Number: W81XWH-08-1-0784

TITLE: The presence of Nonnuclear Estrogen Receptor-alpha in Breast Cancer and Its Prognostic/Predictive Value

PRINCIPAL INVESTIGATOR: Allison Welsh, PhD Candidate

CONTRACTING ORGANIZATION: Yale University, New Haven CT 06510

REPORT DATE: October 2010

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel
               Command
               Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

     X Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (From - To)
01-10-2010	Annual Summary	1 Oct 2009 - 30 Sep 2010

4. TITLE AND SUBTITLE

The presence of non-nuclear Estrogen Receptor-alpha in breasat cancer and its prognostic/predictive value

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-08-1-0784

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Allison Welsh

allison.welsh@yale.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Yale University
New Haven, CT 06510

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

US Army Medical Research and
Matriel Command
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT My research has shown that while we can detect non-nuclear ER reproducibly in a small number of breast cancer specimens, its low incidence implies a minimal prognostic or predictive value, and often poor antibody validation is a confounding factor. While negative, we feel these results are important to share with the scientific and clinical world. These roadblocks have also provided an impetus to develop a quantified and standardized assay to measure nuclear ER itself, and to use this assay in assessing the level and significance of ER misclassification in breast cancer patients today. This has allowed us to provide insight into the current causes of false-negative ER classification, with two especially important and clinically-relevant conclusions that have been a main focus of this year's work: 1) current problems with misclassification appear due to variability in threshold intensity of DAB stain, rather than variability in %-positive cells, and thus new ASCO/CAP guidelines in the future should address this problem. And 2) that SP1 appears to be a potentially more sensitive antibody than 1D5 (showing higher signal-to-noise) and when used clinically, appears to reduce the false-negative rate.

15. SUBJECT TERMS
Estrogen Receptor, non-genomic, standardization, false-negative, ER-beta, ER-alpha-36

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	52	**19b. TELEPHONE NUMBER** (include area code)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

INTRODUCTION:

The ultimate goal of this research is to develop one (or a variety of) assay(s) that can provide superior value in predicting response to Tamoxifen. This goal was motivated by the main biologic concern that the current standard for measuring Estrogen Receptor (ER) has some inherent flaws, which can be distilled into two main problems: 1) the current method for measuring nuclear ER with immunohistochemistry (IHC) is highly subjective and not standardized from lab-to-lab or institution-to-institution, and 2) we only measure nuclear ER expression, despite the fact that it is somewhat widely accepted amongst scientists that ER can function non-genomically as well. This non-genomic signaling has been shown to underlie Tamoxifen resistance in many preclinical models (1-5), and can involve full-length receptor or shorter isoforms (6-9), as well as cross-talk with other GFRs (10-12) and cytoplasmic kinase pathways (13-14). Therefore, the aim of this research is primarily to improve the way we measure nuclear ER itself (by developing a quantitative and standardized method), and secondarily to develop an assay to detect non-genomic signaling. This second aim involved first simply trying to detect non-nuclear ER in actual clinical samples, and then involved efforts to develop an assay (or assays) that could measure different aspects of this non-genomic signaling.

I chose to focus my *initial proposal* on these second set of aims, non-nuclear ER and its cross-talk with Src (as you can see reflected in the original Statement of Work below), since data in the literature suggests this is a major component of non-genomic signaling (1, 12, 15-18). However, I have simultaneously been able to develop an assay for standardizing measurement of nuclear ER, which has now, in part, been adopted for actual clinical use. Finally, this work has allowed me to examine the level of, and causes for, false-negative ER classification in current clinical practice.

BODY:

The original statement of work was the following:

Build an assay to quantitatively assess the activity of non-genomic ERα signaling in breast cancer

Task 1 Construct cell line models for genomic and non-genomic pathways
Methods: Culture MCF7 (genomic model), MCF-7/HER2-18 cells (non-genomic model); cell stimulation (E2, EGF, IGF-1, tamoxifen, EDC) and IB for ERα (in non-nuclear fraction of lysate), pERα, pHER2 (all within minutes); and ERE-gene reporter assay (within hours)
Timeline: Months 1-7
Outcomes/Deliverables: A cell line model displaying high levels of non-genomic ERα signaling (MCF7/HER2-18), as well as one displaying low levels (but high genomic signaling) as a negative control.

Task 2 Validate and develop antibodies to best distinguish activity of the non-genomic pathway
Methods: Selected antibodies, both cell lines: Subcellular fractionation and subsequent IB for ERα; Immunoflouresence (IF) on coverslips (for ERα, Src, active Src); image capture with DeltaVision microscope; co-IP (ERα/Src); IB for pMAPK; reporter assays (ERE, ERK genes); siRNA to ERα, Src; cell line array construction; IF on cell line array (for ERα, Src, active Src); Image capture with PM-200 fluorescence microscope, AQUA analysis
Timeline: Months 4-20
Outcomes/Deliverables: A quantitative non-genomic ERα pathway assay: validated antibodies and IF readouts of non-genomic signaling measured with AQUA

Determine the prognostic and predictive value of the non-genomic ERα pathway assay in breast cancer patients

Task 3 Development and optimization of antibodies for use on human tissue microarrays (TMAs)

Methods: anti-ERα, anti-Src IF: breast test array (antibody titer), ERα boutique array; Image capture with PM-2000 fluorescence microscope; AQUA; analysis of score frequency distribution & linear regression for reproducibility

Timeline: Months 17-22

Outcomes/Deliverables: ERα, Src, and "active" Src antibodies optimal for IF on full TMAs

Task 4 Assessment of the prognostic and predictive value of the non-genomic ERα pathway assay using a large patient cohort

Methods: IHC using anti-ERα, anti-Src, anti-"active Src" on full cohort TMA (majority with long-term follow up); Image capture with PM-200 flourescence microscope; AQUA analysis; score frequency distribution, cut point analysis; Clinical data retrieval; Generation of Kaplan-Meier survival curves; Univariate and multivariate analyses; IF using anti-ERα, anti-Src, anti-"active Src" on special cohorts (300 patient Yale cohort and Swedish cohort, both with tamoxifen treatment); Image capture with PM-200 fluorescence microscope, AQUA analysis, Clinical data retrieval; Generation of Kaplan-Meier survival curves, Univariate and multivariate analyses

Timeline: Months 22-36

Outcomes/Deliverables: Determination of the prognostic and/or predictive value of non-genomic ERα pathway assay for breast cancer patients

As I explained in detail in my last progress report (see Oct 2009), I had decided over a year ago to put **Task 1** on hold (and potentially move beyond it altogether), when we realized how unreproducible and variable cell line models were, and how unable they were to faithfully represent what we observe in actual patient tumors. Especially in the case of non-nuclear ER, extensive research in cell line models has already been published, but the real challenge has been proving these same functions are present in actual human tissue.

In my last progress report (see Oct 2009), I documented work on **Tasks 2-4,** including the following topics:

- **Validation of full-length ER antibodies (multiple epitopes) and development for use on TMAs (Tasks 2 & 3)**
- **Development of an assay to quantify non-nuclear (cytoplasmic) ER in patient tissue & assessment of prognostic value (Task 3 & 4)**
- **Development of an assay to reproducibly quantify nuclear ER in patient tissue & assessment of current ER misclassification rate**
- **Validation of antibodies to Src and pER and development for use on TMAs (Tasks 2 & 3)**

Therefore over the course of the past year, I have continued work on these areas & proceeded further in areas where I found promising results. The summary of my work over the past year is organized under the following aims:

Assess non-genomic ER pathway
      1) Cytoplasmic ER: assess presence & significance in clinical samples (Tasks 2-4)
      2) Develop assay to assess significance of non-genomic ER proteins (ERβ, ER36) (Tasks 2-4)

Improve assessment of nuclear ER
> **3) Develop Q-IF Assay: quantitative & standardized assay for nuclear ERα**
> **4) Determine level of ER misclassification (discordance) due to lab-to-lab variability in DAB staining in current US practice**

1) Cytoplasmic ER: assess presence & significance in clinical samples (Tasks 2-4)
The *hypothesis* for this aspect of my project has been that patients with high levels of cytoplasmic ERα will respond worse to endocrine therapies than patients with lower levels, based on the biology of non-nuclear ER function and cross-talk with growth-factor receptor pathways currently published (1-16). However, a *main problem* we have faced is that all evidence of cytoplasmic ER thus far has been shown in cell line or mouse models (1, 2, 5-8, 10, 12-16). There has been no concrete evidence of existence in clinical cases.

As I reported last year, we have shown that multiple antibodies to different epitopes of ERα are highly specific and reproducible (**Fig 1**). We have also shown that when full-length ER is localized to the cytoplasm (by an engineered mutation in the nuclear localization sequence, NLS) we are still able to detect it with this panel of antibodies (**Fig 2**). This model employed GFP-tagged wild-type ER or cyto-ER (mutated NLS), which were overexpressed in MCF-7 cells (a cell line which also harbors endogenous, non-GFP-tagged, ER). In this model, only a mutation in the NLS was necessary to confer cytoplasmic localization of ER, therefore it does not recapitulate the variety of other possible forms of ER that could be present outside the nucleus (alternative isoforms, post-translational modifications).

When we looked for cytoplasmic localization of ER in actual clinical samples, by immunofluoresence (IF), we were able to detect it using multiple antibodies from the panel (**Fig 3**). This evidence suggests that the cytoplasmic ER we observe is not an epitope-specific artifact, and furthermore that at least a portion of it is full-length receptor (or a form of ER with both the N- and C-terminus intact). However, ultimately, after examining a number of different retrospective breast cancer cohorts, we found the incidence of cytoplasmic ER to be very low overall (**Table 1**). It ranged from 1-3% on average, and was further complicated by the fact that we sometimes observed cytoplasmic ER in conjunction with strong nuclear staining, raising the question as to whether, if it is not an artifact, it is more important to measure total cytoplasmic levels or the ratio of cytoplasmic to nuclear staining. Only one cohort (B14) showed a relatively high percentage of cytoplasmic cases (10%), and this was part of a collaborative study whose terms we agreed to prospectively, and thus we were unable to retrospectively perform experimental analyses on these cytoplasmic cases. Furthermore, this is an old cohort, and the methods of fixation were noted to be extremely variable. We have evidence to suggest that ER protein levels decrease as a function of time-to-fixation (17), which raises the question whether the cytoplasmic localization of ER may correlate with different fixation methods or ischemic times.

Despite the unlikelihood of developing a prognostic/predictive assay, we did attempt to perform an exploratory analysis to determine the identity of the cytoplasmic reactivity we saw with the ER antibodies (**Fig 3**). We hand-picked the small number of clinical cases with visible cytoplasmic staining by IF, pulled their formalin-fixed paraffin-embedded (FFPE) tissue block from our archives, and took a sample core. We then performed RNA extraction on each sample, assessed the concentration & purity using the Nanodrop technology, and performed RT followed by PCR for ER as well as β-actin. We were able to perform successful RT-PCR on RNA prepared fresh

from cell lines, however, the RNA was too degraded in our FFPE samples to get any intact PCR product (even the 100kB β-actin).

Lastly, we found that the process of antibody validation is critical to assessment of non-nuclear ER in clinical specimens. We tested a protocol developed by a collaborator in the field, who claimed to have found cytoplasmic ER in clinical specimens, and found the antibody (MC20) to be responsible for the observed cytoplasmic staining. When using western blot analysis of a cell line panel at short exposure, MC20 appears to give a specific band at the expected size of 66kD (**Fig 4a, top panel**). However, upon a longer 1min exposure, the MC20 antibody reveals multiple immunoreactive bands in all cell lines and at various sizes, in stark contrast to the specific band in the three known-positive cell lines as observed with SP1 antibody (**Fig 4a, bottom panels**). Furthermore, when comparing both antibodies using IF on the cell line panel, again we see lack of the specificity with MC20 that is present with SP1 (**Fig 4b**). Figure 5 shows representative IF images of staining with MC20 and SP1 (**Fig 5a**), showing how MC20 could appear to be cytoplasmic ER. However, when both antibodies are stained on a control set of MCF-7 cells overexpressing tet-inducible ER, we see that increasing amounts of doxycycline increase the specific nuclear staining with SP1, but have no affect on the non-specific staining seen with MC20 (**Fig 5b**).

We have therefore come to the conclusion that, given the limitations/caveats we have outlined, the overall incidence of cytoplasmic ER is too low to be of prognostic or predictive value as an assay. All of this data is in the process of being assembled in a manuscript which we plan to submit to *Breast Cancer Research*. Future studies, however, will look into the relationship between cytoplasmic expression and ischemic/fixation time, as well as the possibility of alternatively spliced isoforms of ER outside the nucleus. Additional projects could also look into the presence of non-nuclear ER in neoadjuvant specimens (preclinical research suggests Tamoxifen treatment may induce a higher degree of cytoplasmic localization).

2) Develop assay to assess significance of non-genomic ER proteins (ERβ, ER36) (Tasks 2-4)
I have also been working on developing an assay to assess non-genomic ER function, by measuring other proteins reported to be involved in this signalling. Key players whose expression was suggested to be of prognostic/predictive value on their own were the isoforms of ERβ (encoded by a separate gene than ERα) and ER36 (a short isoform of ERα, alternatively spliced with a unique 27aa C-terminal sequence, and proposed to be primarily membranous/cytoplasmic) (18-22).

Much of the work in our lab has focused on antibody validation as a critical component of any studies which use them, and this has most often been the biggest obstacle in many of our projects. We have developed an extensive protocol which we use in the lab and I contributed to two publications on this topic, which I am not appending, but are listed in the references (23-24).

I began by validating antibodies to ERβ1, ERβ2, and ERβ5, but none of them were usable for western blot analysis, one of our standard validation procedures. In cell lines which were engineered to overexpress a tet-inducible ERβ1 or ERβ2, we did not observe an increase in immunoreactivity by IF upon induction (2ug/ml doxy) for either antibody (**Fig 6, right panels**), and upon RNA silencing of total ERβ, at best observed a modest decrease in immunoreactivity with the ERβ2 antibody (**Fig 6, bottom panels**). Furthermore, both antibodies showed both nuclear and cytoplasmic staining in FFPE clinical cases (**Fig 7a**), which had a specificity that was difficult to validate. Both also showed poor reproducibility for total staining on duplicate cores (shown for ERβ1 in **Fig7b**, similar results found with ERβ2). Because of these reasons, and

discussions with collaborators who had found similar problems, we decided not to proceed with these antibodies, but are working in collaboration with Cell Signaling to produce specific and reproducible antibodies to the ERβ isoforms that can be used in the near future.

A similar problem was encountered with ER36, however in this case, the only currently available antibody was not commercial, but developed by Dr. Wang who initially discovered and cloned ER36. He sent us an aliquot of his antibody, but we could not reproduce his data. We have been working closely with Cell Signaling on this project, and the development of this antibody (specific to the 27aa sequence at the C-terminus of ER36) has been one of their highest priorities, with 30 rabbits and various immunogen designs currently in late stages of development. We have also been working in collaboration with Rachel Schiff at Baylor College of Medicine, who has been producing the ideal cell line models in which to validate this antibody once we receive it. They have currently developed transient transfections of a FLAG-tagged ER36 in Hek293 and MCF7 cells, and are fixing these in formalin and embedding them in paraffin, so we can construct a control TMA from cores. They have had more difficulty producing lines stably transfected with ER36 (suspect it is potentially lethal in cell line models), but are continuing work on this front as well. As soon as the antibody is ready, we will have an ideal system to rapidly validate it, and proceed to development of an IF-based assay on TMAs.

3) Develop Q-IF Assay: quantitative & standardized assay for nuclear ERα
Last year I reported on the development of a quantitative & standardized assay to measure nuclear ER (see Oct 2009). This project was inspired by the inherent subjectivity involved in the current IHC test for ER, and the problems with false-negative classification of patients that has been reported in the literature recently (25-27). We used it to look at the level & significance of discordance in ER status on two retrospective cohorts here at Yale. I had submitted the paper to *JNCI* in Jan 2010, and after all revisions, good feedback & signing the final forms, it was rejected suddenly at the end of June.

Concordantly, the new ASCO/CAP guidelines for ER testing had been released on June 1 (28), which lowered the threshold for what is considered ER-positive from 10%-positive nuclei to 1%-positive nuclei. This change was designed to address the false-negative rate, and therefore would presumably help fix the problem I had raised in the submitted paper. However, our data strongly suggests that the discordance in ER status is due to *intensity* of nuclear staining for ER, rather than the *percentage* of positive cells. In other words, the problem is what we consider to be a "positive" nuclei, rather than how many there are (or at least in addition to how many there are), but the guidelines only define positive as "any immunoreactivity". We therefore re-analyzed the retrospective cohort we have access to (YTMA 49) with the new guidelines, and found the same exact results, that is, the level discordance doesn't change as a result of the switch from 10% to 1%. We subsequently re-wrote the paper, added this new data, and re-submitted to *Journal of Clinical Oncology* on Oct 1. It is now under final stages of review, and we expect it will be accepted and published soon. Because the majority of the figures were shown in my last progress report, I am appending the entire paper at the end on this report, titled "Standardization of Estrogen Receptor Measurement in Breast Cancer Suggests False Negative Results are a Function of Threshold Intensity Rather than Percentage of Positive Cells".

In terms of clinical implications, much of this data on quantification and standardization was actually translated (and is cited in their marketing material) into an ER testing platform by Genoptix, Inc (Carlsbad, CA), just released at the San Antonio Breast Cancer Symposium Dec 8-12, and is now available to clinicians across the country at the same cost as traditional IHC. Already, over 30 patients have been tested, or re-tested, for ER status using the technology and 2 have already switched from ER negative to positive.

4) Determine level of ER misclassification (discordance) due to lab-to-lab variability in DAB staining in current US practice

As a follow-up to these studies, we then asked: what is the cause of the discordance in intensity? Is Q-IF more sensitive than IHC? Is it due to variation in DAB from lab-to-lab? To determine this, the Quantitative Immunofluorescence (QIF) assay using the AQUA method was performed on our control array (the Index TMA), containing 40 patient controls, and analyzed on a case-by-case basis, comparing QIF to IHC done by routine protocol (or without the hematoxylin counterstain) in two labs. We also performed a more in-depth analysis on our large retrospective Yale cohort (YTMA 49) in order to further compare the variability in threshold for positivity. YTMA 49 was stained by routine IHC in 4 labs (3 clinical, 1 research; 3 used Dako 1D5 antibody system, 1 used SP1 Ventana), followed by analysis by three individuals (myself, as well as two board-certified pathologists- MH and DLR) who scored for both intensity (0-3) and %-positive (0-100) cells. IHC scores for each case were binarized into ER positive/negative using both the old (10%) or the new (1%) threshold guidelines, and ER status was then compared lab vs. lab and 10% vs. 1%. ER status in YTMA 49 was also determined twice by the QIF assay (once using 1D5 antibody and once with SP1) in order to compare discordance in ER status due to method (IHC vs. QIF) as well as antibody choice (1D5 vs. SP1).

In the Index TMA, 19 of 31 scoreable cases were ER positive by QIF. By routine IHC, three of these had discordant ER status (1/3 negative in Lab1, 3/3 negative in Lab2). However, when IHC was performed without hematoxylin, low levels of ER were visible above background in all 3 cases (**Fig 8**). This suggested that subtle levels of ER are detectable by QIF, but not by routine IHC tests that include a hematoxylin counterstain. On YTMA 49, we found 10-32% of cases to have discordant ER status depending on the Lab where IHC was performed (**Table 2**). However, as expected, we found discordance levels did not significantly change when using a 10% or 1% threshold (**Table 2**). When we examined only the discordant cases, and looked at the scores for %-positive, we found them evenly distributed across a range from 5% -100%, with the majority well above the 1-10% threshold, providing further evidence that discordance isn't due to a discrepancy in %-positive threshold (**Fig 9**). Examples of two discordant cases are shown in **Figure 10**.

We then performed Kaplan-Meier disease-specific survival analysis of all subgroups of patients, discordant lab-to-lab (**Fig 11**) or QIF-to-IHC (**Fig 12**). These analyses revealed that discordant cases showed survival behavior similar to double positives (both assays ER positive), suggesting they are actually false-negatives, and thus potentially under-treated.

Lastly, we examined the level of discordance due to antibody choice. 1D5 (and the Dako system) are the most common standard used clinically, but more recently SP1 (commercially available from Ventana) has been used as well, and some published data (29-30), as well as our own findings, suggests SP1 may have higher signal. Since one lab used the Ventana system, we could compare discordance in IHC due to SP1 vs. 1D5 and found that to be 18%. When examining the cell line panel as well as the 40 patient controls on our Index array, we found both antibodies to have the same threshold for positivity (same cases were considered positive and negative), however we saw a much greater signal to background ratio with SP1 (**Fig 13, A-D**). In other words, the difference between the highest negative case and the lowest positive case was much more pronounced with SP1, and even visible by eye (**Fig 10, four right panels**). When we examined the full cohort (YTMA 49) by QIF with SP1 versus 1D5, we did find 8.8% of cases to have discordant ER status (**Fig 13 E**), with almost all of these positive by SP1 but negative by 1D5. Lastly, we performed Kaplan-Meier survival analysis of these cases, and found these discordant cases (SP1+/1D5-), to show outcome behavior similar to the double-positives (**Fig 13**

9

F), suggesting that use of 1D5 results in an increased level of false-negative cases. Caveats for this study include the fact that it was done on TMAs instead of whole sections, but we have been able to reproduce the level of discordance observed on a second retrospective cohort (however it is too recent to have follow-up information). We are in the process of putting together this data for publication and plan to submit it to *Modern Pathology*.

KEY RESEARCH & TRAINING ACCOMPLISHMENTS:

1) Validated that four different monoclonal antibodies against multiple epitopes of full-length ER (1D5 the clinical gold-standard, as well as F10, SP1, and 60c) are specific and generally equivalent in their detection of nuclear ER by western blot, by IF on cell lines, and by IF on tissue microarrays.
2) Validated that all four monoclonal antibodies above can detect cytoplasmic ER in cell lines and clinical cases
3) Found that incidence of cytoplasmic ER in untreated clinical cases was too low to be of use as a prognostic/predictive marker alone
4) Helped develop a protocol/schematic for successful antibody validation
5) Collaborated with Cell Signaling and Rachel Schiff Lab at Baylor College of Medicine to develop monoclonal antibody to ER36 and prepare cell line models for effective validation
6) Developed an assay to standardize quantification of nuclear ER in patient tissue using an Index of Control Cases
7) Witnessed translation of developed assay into a commercially available technology through Genoptix, Inc.
8) Found a 10-30% level of discordance in ER status between clinical labs using traditional DAB-staining for IHC analysis.
9) Showed the level of discordance in ER status appears due to threshold (what is considered a "positive" nuclei) rather than the %-positive cells
10) Used standardized AQUA-based ER assay to show that QIF methods can detect subtle levels of ER that are not detectable by routine IHC tests that include a hematoxylin counterstain
11) Showed that a significant degree of discordance in ER status (9-18%) is due to antibody choice, where SP1 shows higher signal to noise (potentially more sensitive).

REPORTABLE OUTCOMES:
Manuscripts – First author:
1) *Standardization of Estrogen Receptor Measurement in Breast Cancer Suggests False Negative Results are a Function of Threshold Intensity Rather than Percentage of Positive Cells.* Welsh AW, Moeder C, Alarid E, Haffty B, Rimm DL. Submitted to JCO, currently under final revisions.

Manuscripts – second author:
1) Anagnostou VK, Welsh AW, Giltnane JM, Siddiqui S, Liceaga C, Gustavson M, Syrigos KN, Reiter JL, & Rimm DL (2010). Analytic variability in immunohistochemistry biomarker studies. *Cancer Epidemiol, Biomarkers Prev* **19**: 982 – 991
2) Bordeaux J, Welsh A, Agarwal S, Killiam E, Baquero M, Hanna J, Anagnostou V, & Rimm D (2010). Antibody validation. *Biotechniques* **48**: 197-209.

Abstracts & Poster presentations:
1) Poster Discussion: 33rd Annual San Antonio Breast Cancer Symposium, December 8-12, 2010. *Causes for false-negative Estrogen Receptor (ER) classification in breast cancer.* Allison Welsh, Malini Harigopal, and David L. Rimm.
2) Abstract submission: 100th annual USCAP meeting Feb 26-Mar 4, 2011. *Discordance for Estrogen Recpetor (ER) Status Between Labs is Still Very High, Despite ASCO/CAP Guidelines*. Allison Welsh, Malini Harigopal, and David L. Rimm.

3) Abstract & Poster Presentation: 2009 American Society of Clinical Oncology (ASCO) Annual Meeting. *Evaluation of the false-negative rate of standardized and quantitative measurement of estrogen receptor (ER) in tissue using AQUA technology (#567 Booth #2679)*

4) Abstract & Poster Presentation: 32nd Annual San Antonio Breast Cancer Symposium, December 9-13, 2009. *Development of a quantitative and standardized assay to measure ER protein concentration in breast cancer tissue & improve current patient misclassification (#4068)*

Talks/Presentations:
1) Presentation at Cambridge Healthtech Institute's second annual Science of Biobaking conference, Dec 6-8, 2010, Providence RI: *Extrinsic & Intrinsic Controls for Measurement of Protein Analyte Concentrations in Tissue Slides.*

2) Yale University Department of Pathology Research in Progress talk. One each year: March 2010, March 2009.

Degrees Obtained:

Expected completion of PhD in Pathology from Yale University School of Medicine, March 2011.

CONCLUSION:

In conclusion, much of my work on the functional role of cytoplasmic ER in breast cancer has revealed a minimal prognostic or predictive value. However, the work itself has proven an invaluable learning experience and led to submission of a manuscript (in progress) on these results, which, while negative, we feel are important to share with the scientific and clinical world. Furthermore, these findings have allowed me to focus on a more basic clinical problem regarding measurement of ER: the problem of subjectivity and variability in assessment of nuclear ER itself. My work to date has allowed me to develop a quantified and standardized assay to measure nuclear ER, and to use this assay in assessing the level and significance of ER misclassification in breast cancer patients today. This has allowed us to provide insight into the current causes of false-negative ER classification, with two especially important and clinically-relevant conclusions: 1) that current problems with misclassification appear due to variability in threshold intensity of DAB stain, rather than variability in %-positive cells, and thus new ASCO/CAP guidelines in the future must address this problem. And 2) that SP1 appears to be a potentially more sensitive antibody than 1D5 (showing higher signal-to-noise) and when used clinically, appears to reduce the false-negative rate.

While these studies have their own limitations (use of TMAs instead of whole sections, use of cohorts with only prognostic instead of predictive information), they have still led to publications, abstracts and talks that I feel privileged, as a graduate student, to have experienced this early in my career. Furthermore, I have tangibly felt their clinical impact with the development of an AQUA-based ER testing platform by Genoptix, Inc (whose marketing material cites this research as their first reference). Again, as a graduate student, this experience has been incredibly humbling and inspiring, and it could not have been possible without the support of this funding.

REFERENCES:

1. Hiscox S, Morgan L, Green TP, Barrow D, Gee J & Nicholson RI (2006). Elevated Src activity promotes cellular invasion and motility in tamoxifen resistant breast cancer cells. *Breast Canc Res & Treat* **97:** 263-274.
2. Fan P, Wang J, Santen RJ & Yue W (2007). Long-term treatment with tamoxifen facilitates translocation of estrogen receptor alpha out of the nucleus and enhances its interaction with EGFR in MCF-7 breast cancer cells. *Canc Res* **67:** 1352-1360.
3. Massarweh S & Schiff R (2007). Unraveling the mechanisms of endocrine resistance in breast cancer: New therapeutic opportunities. *Clin Cancer Res* **13:** 1950-1954.
4. Gururaj AE, Rayala SK, Vadlamudi RK & Kumar R (2006). Novel mechanisms of resistance to endocrine therapy: Genomic and nongenomic considerations. *Clin Canc Res* **12 (3 Suppl):** s1001-s1007.
5. Wiseman LR, Johnson MD, Wakeling AE, Lykkesfeldt AE, May FE & Westley BR (1993). Type I IGF receptor and acquired tamoxifen resistance in oestrogen-responsive human breast cancer cells. *Eur J Cancer* **29A:** 2256-64.
6. Li L, Haynes MP & Bender JR (2003). Plasma membrane localization and function of the estrogen receptor alpha variant (ER46) in human endothelial cells. *Proc Natl Acad Sci U S A* **100:** 4807-12.
7. Wang Z, Zhang X, Shen P, Loggie BW, Chang Y & Deuel TF (2006). A variant of estrogen receptor-{alpha}, hER-{alpha}36: transduction of estrogen- and antiestrogen-dependent membrane-initiated mitogenic signaling. *Proc Natl Acad Sci U S A* **103:** 9063-8.
8. Acconcia F, Ascenzi P, Bodeci A, Spisni E, Tomasi V, Trentalance A, Visca P & Marino M (2005). Palmitoylation-dependent estrogen receptor-alpha membrane localization regulation by 17 beta-estradiol. *Mol Biol Cell* **16:** 231-237.
9. Boonyaratanakornkit V & Edwards DP (2007). Receptor mechanisms mediating non-genomic actions of sex steroids. *Semin Reprod Med* **25:** 139-53.
10. Schiff R, Massarweh SA, Shou J, Bharwani L, Arpino G, Rimawi M & Osborne CK (2005). Advanced concepts in estrogen receptor biology and breast cancer endocrine resistance: implicated role of growth factor signaling and estrogen receptor coregulators. *Cancer Chemother Pharmacol* **56 (Suppl 1):** s10-s20.
11. Baserga R, Peruzzi F & Reiss K (2003). The IGF-1 receptor in cancer biology. *Int J Cancer* **107:** 873-7.
12. Hitosugi T, Sasaki K, Sato M, Suzuki Y & Umezawa Y (2007). Epidermal growth factor directs sex-specific steroid signaling through Src activation. *J Biol Chem* **282:** 10697-10706.
13. Greger JG, Fursov N, Cooch N, McLarney S, Freedman LP, Edwards DP & Cheskis BJ (2007). Phosphorylation of MNAR promotes estrogen activation of phosphatidylinositol 3-kinase. *Mol Cell Biol* **27:** 1904-1913.
14. Wong CW, McNally C, Nickbarg E, Komm BS & Cheskis BJ (2002). Estrogen receptor-interacting protein that modulates its nongenomic activity-crosstalk with Src/Erk phosphorylation cascade. *PNAS* **99:** 14783-14788.
15. Yang Z, Barnes CJ, Kumar R (2004). Human epidermal growth factor receptor 2 status modulates subcellular localization of and interaction with estrogen receptor alpha in breast cancer cells. *Clin Canc Res* **11**: 3621-3628.
16. Shupnik MA (2004). Crosstalk between steroid receptors and the c-Src-receptor tyrosine kinase pathways: implications for cell proliferation. *Oncogene* **23:** 7979-7989.
17. Bai Y, Tolles J, Cheng H, Siddiqui S, Gopinath A, Pectasides E, Camp RL, Rimm DL, Molinaro AM (2010). Quantitative Assessment Shows Loss of Antigenic Epitopes as a Function of Time to Formalin Fixation. In submission.
18. Speirs V (2008). The evolving role of oestrogen receptor beta in clinical breast cancer. *Breast Cancer Res* **10**: 111.
19. Shaaban AM, Green AR, Karthik S, Alizadeh Y, Hughes TA, Harkins L, Ellis IO, Robertson JF, Paish EC, Saunders PT, Groome NP, Speirs V. Nuclear and cytoplasmic expression of ERbeta1, ERbeta2, and ERbeta5 identifies distinct prognostic outcome for breast cancer patients. *Clin Canc Res* **16**: 5228 – 5235.

20. Shi L, Dong B, Li Z, Lu Y, Ouyang T, Li J, Wang T, Fan Z, Fan T, Lin B, Wang Z, & Xie Y (2009). Expression of ERα36, a novel variant of estrogen receptor α, and resistance to tamoxifen treatment in breast cancer. *J Clin Oncol* **27**: 3423 – 3429.

21. Lee LM, Cao J, Deng H, Chen P, Gatalica Z, Wang ZY (2008). ER-alpha36, a novel variant of ER-alpha, is expressed in ER-positive and -negative human breast carcinomas. *Anticancer Res* **28**: 479 – 493.

22. Wang Z, Zhang X, Shen P, Loggie BW, Chang Y, Deuel TF (2006). A variant of estrogen receptor-{alpha}, hER-{alpha}36: transduction of estrogen- and antiestrogen-dependent membrane-initiated mitogenic signaling. *Proc Natl Acad Sci* **103**: 9063 – 9068.

23. Anagnostou VK, Welsh AW, Giltnane JM, Siddiqui S, Liceaga C, Gustavson M, Syrigos KN, Reiter JL, & Rimm DL (2010). Analytic variability in immunohistochemistry biomarker studies. *Cancer Epidemiol, Biomarkers Prev* **19**: 982 – 991.

24. Bordeaux J, Welsh A, Agarwal S, Killiam E, Baquero M, Hanna J, Anagnostou V, & Rimm D (2010). Antibody validation. *Biotechniques* **48**: 197-209.

25. Hede K (2008). Breast cancer testing scandal shines spotlight on black box of clinical laboratory testing. *JNCI* **100**: 836-7.

26. Allison KH (2008). Estrogen receptor expression in breast cancer: we cannot ignore the shades of gray. *Am J Clin Pathol* **130**: 853-4.

27. Allred DC (2008). Commentary: hormone receptor testing in breast cancer: a distress signal from Canada. *Oncologist* **13**: 1134-1136.

28. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, Fitzgibbons PL, Francis G, Goldstein NS, Hayes M, Hicks DG, Lester S, Love R, Mangu PB, McShane L, Miller K, Osborne CK, Paik S, Perlmutter J, Rhodes A, Sasano H, Schwartz JN, Sweep FC, Taube S, Torlakovic EE, Valenstein P, Viale G, Visscher D, Wheeler T, Williams RB, Wittliff JL, & Wolff AC (2010). American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Onocol* **28**: 2784-2795.

29. Brock JE, Hornick JL, Richardson AL, Dillon DA, Lester SC (2009). A comparison of estrogen receptor SP1 and 1D5 monoclonal antibodies in routine clinical use reveals similar staining results. *Am J Clin Pathol* **132**: 396-401.

30. Cheang MC, Treaba DO, Speers CH, Olivotto IA, Bajdik CD, Chia SK, Goldstein LC, Gelmon KA, Huntsman D, Gilks CB, Nielsen TO, Gown AM (2006). Immunohistochemical detection using the new rabbit monoclonal antibody SP1 of estrogen receptor in breast cancer is superior to mouse monoclonal antibody 1D5 in predicting survival. *J Clin Oncol* **36**: 5637-5644.
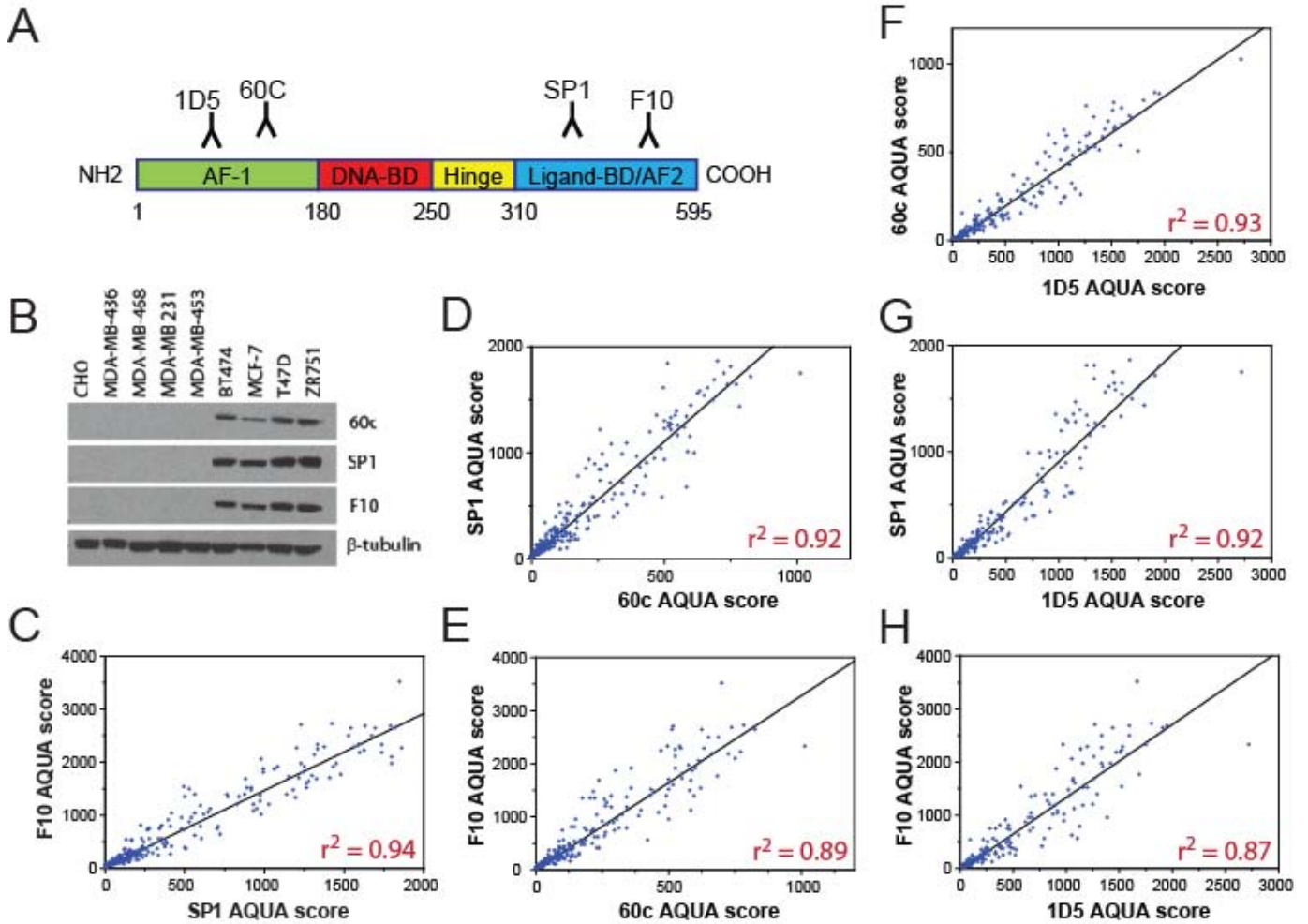
Figure 1. Multiple antibodies to different epitopes of ERα are highly specific and reproducible. A) Schematic of four monoclonal antibodies to ERα and their mapping to epitopes of ERα. **B)** Western blot analysis of ER in breast cancer cell line panel (positive controls BT474, MCF7, T47D, ZR751) showing antibody specificity. Antibodies were also used for IF analysis of ER expression (reported as AQUA score) in a retrospective cohort of 650 cases of breast cancer from Yale (YTMA 49, FFPE cases on tissue microarray). **C-H)** Regression between AQUA scores for each antibody, showing high reproducibility.
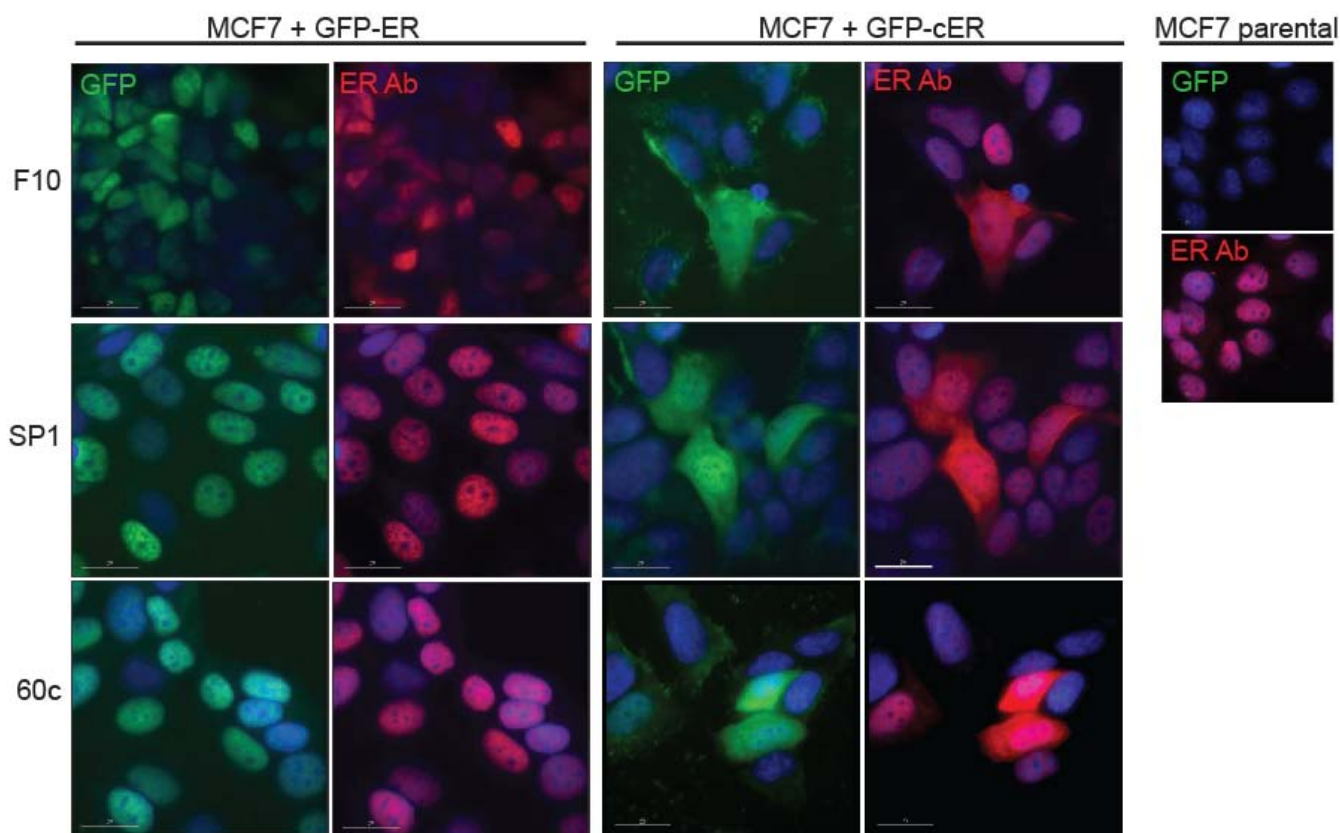
Figure 2. Antibodies to ERα can detect cytoplasmic ERα when it is engineered in cell lines. MCF7 cells were grown and stably transfected with a GFP-tagged wild-type ER (MCF7 + GFP-ER) or a GFP-tagged cytoplasmic ER (cER), which was a deletion mutant lacking its nuclear localization sequence (MCF7 + GFP-cER). These cells were cultured, grown on coverslips, fixed, and stained using immunofluoresence with three antibodies to the N- and C-terminus of ER **(shown in red)**, along with DAPI and GFP. All three antibodies (F10 and SP1 C-terminal, 60c N-terminal) were able to recognize strong, specific nuclear staining for ER (co-localized with GFP) in the GFP-ER cells **(red, left panels)**. All three were also able to recognize strong, specific cytoplasmic staining for ER (also co-localized with GFP) in the GFP-cER cells **(red, right column of panels)**. Endogenous ER **(red)** can be seen in MCF7 controls **(right-most panels)** untransfected with a GFP-tagged construct.
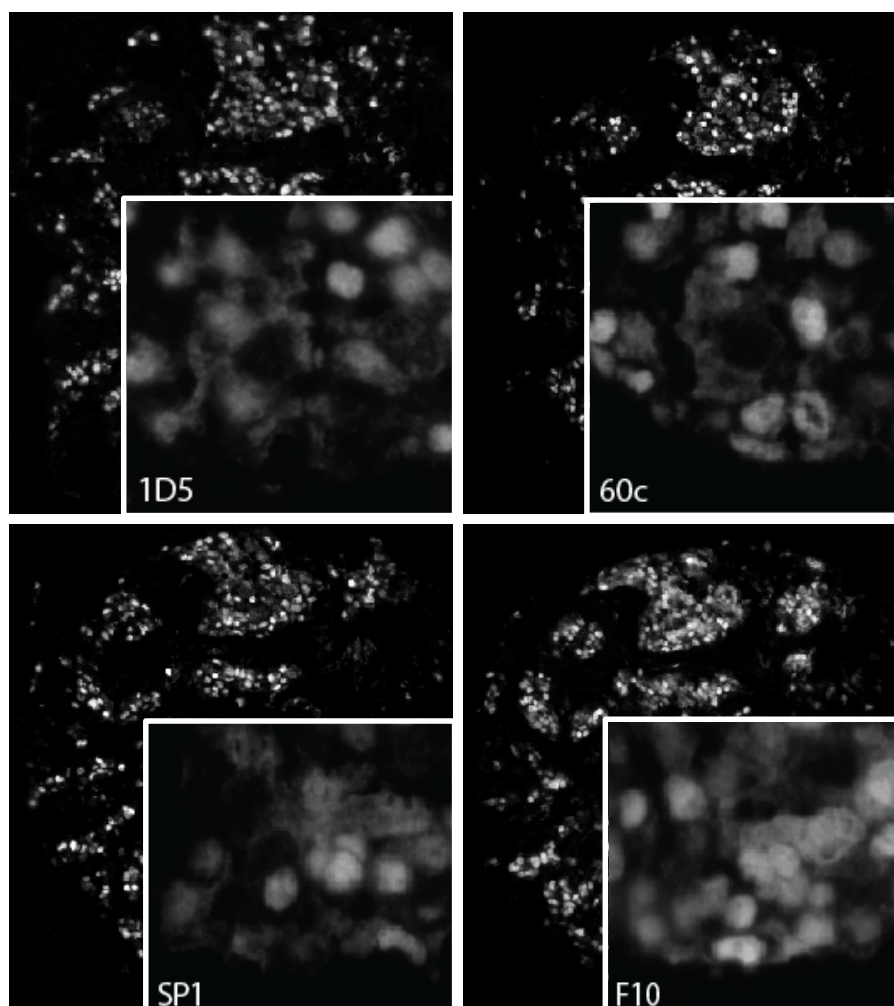
Figure 3. **Detection of cytoplasmic ERα across multiple epitopes in patient samples.** Cytoplasmic ER was detected in FFPE breast cancer specimens present on a retrospective cohort from Yale (YTMA 49). One of the two cases showing strong cytoplasmic localization is shown, revealing cytoplasmic immunoreactivty with all four antibodies.

Cohort	Number of cytoplasmic cases	Number of total cases	Percent (%) of cases with cytoplasmic staining
YTMA 49	4	661	0.6
YTMA 130	4	526	0.7
NSABP B14	60*	657	9.1*
YTMA 128	0	183	0
Richard Love	0	150	0
Total	68	2177	3.1

Table 1. Incidence of cytoplasmic ERα in multiple patient cohorts. Five different retrospective cohorts of breast cancer patients were analyzed on TMAs using IF and AQUA analysis, and cases with specific cytoplasmic staining were hand-counted. Total incidence of cytoplasmic ER was 3.1%. *number of cases were estimated on B14 cohort due to variability of threshold definition for cytoplasmic staining.
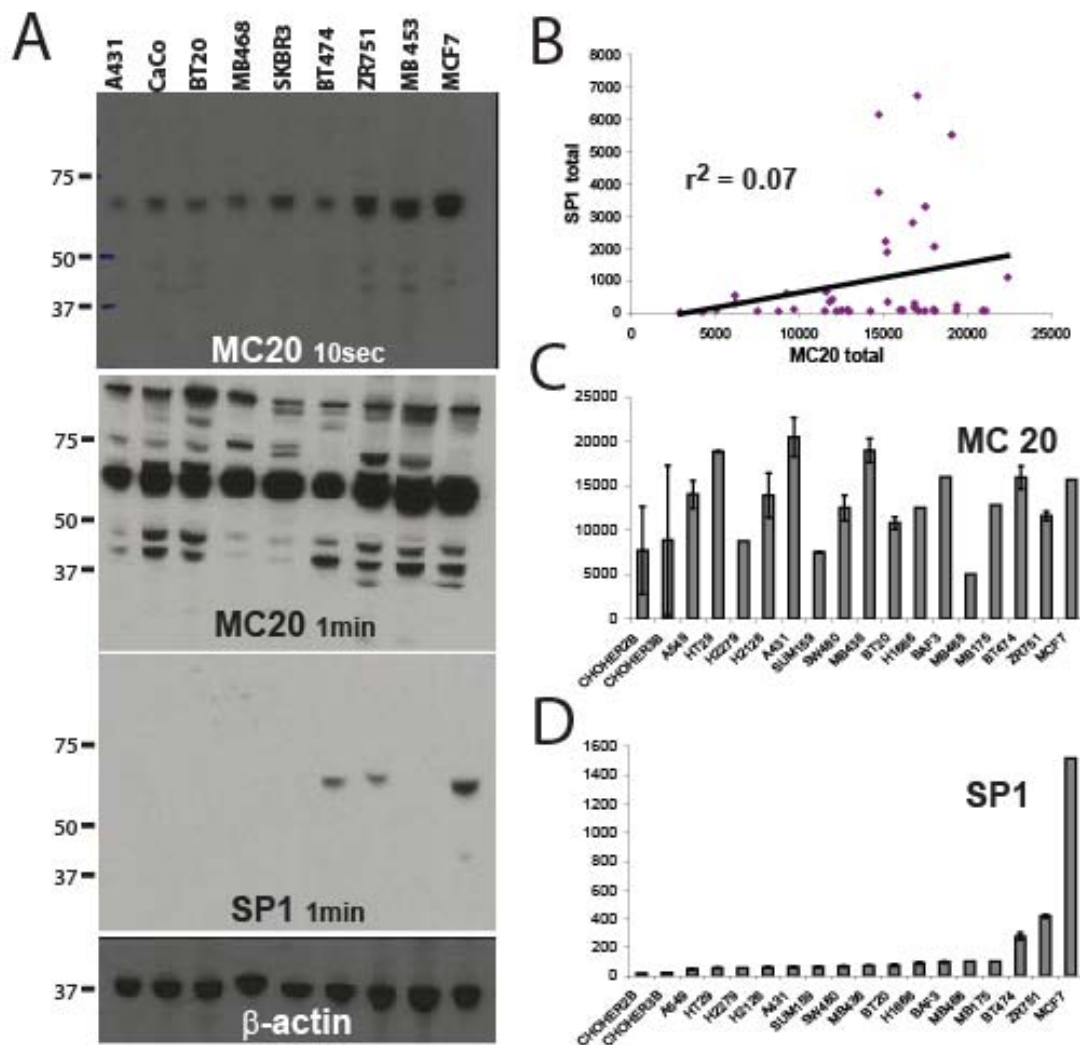
Figure 4. Evidence of cytoplasmic ERα in cell line panel due to invalid antibody. A panel of ATCC breast cancer cell lines were analyzed by western blot using optimal dilutions of the MC20 antibody (rabbit polyclonal, Santa Cruz, which has been reported to detect cytoplasmic ER) or SP1 antibody (rabbit monoclonal, Thermo). The cell lines and a panel of 40 patient controls were also analyzed by IF on TMAs with both antibodies. **A)** Short exposure of MC20 blot could appear to show specific detection of ERα (66kD), but longer exposure of blot shows multiple immunoreactive bands, even in known ER-negative cell lines. SP1, by contract, shows specific reactivity in the three known ER-positive cell lines (BT474, ZR751, MCF7). **B)** Regression analysis of IF AQUA scores for ER in 40 patient controls analyzed with both antibodies showed no correlation ($r^2 = 0.07$). Distribution of If AQUA scores in cell line panel shows non-specific immunoreactivity in all cell lines (including those ER-negative) with MC20 (**panel C**), while SP1 (**panel D**) shows only specific positive AQUA scores for the three cell lines, in agreement with western data. This data was repeated with a second lot of MC20, and the same results were found.
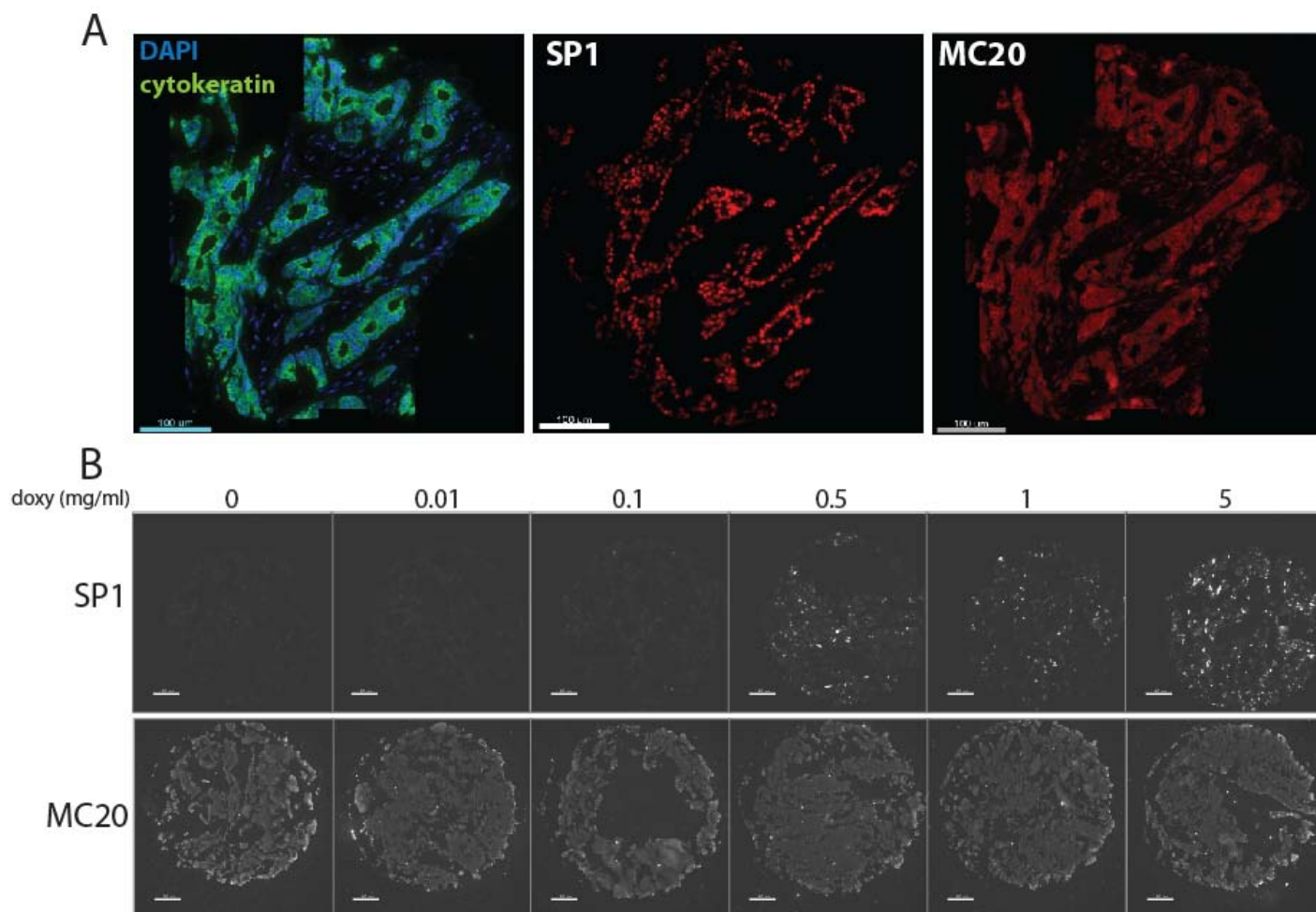
Figure 5. Immunofluorescent evidence of cytoplasmic staining for ERα due to invalid antibody. A panel of 40 patient controls and cell lines were stained for ER using IF techniques with both SP1 (rabbit monoclonal, Thermo) and MC20 (rabbit polyclonal, Santa Cruz, reported to detect cytoplasmic ER) antibodies. Representative IF images are shown in **A** of a patient case, where specific nuclear staining is seen with SP1, but cytoplasmic staining (which could be interpreted as specific) is seen with MC20. Analysis of MCF7 cells with tet-inducible ER overexpression, shows increasing amounts of nuclear reactivity with SP1, in response to increasing amounts of doxycycline **(B, top panels)**, in stark contrast to unchanging levels of "cytoplasmic" staining seen with MC20 **(B, bottom panels)**, proving the non-specificity of the antibody. This data was repeated with a second lot of MC20, and the same results were found.
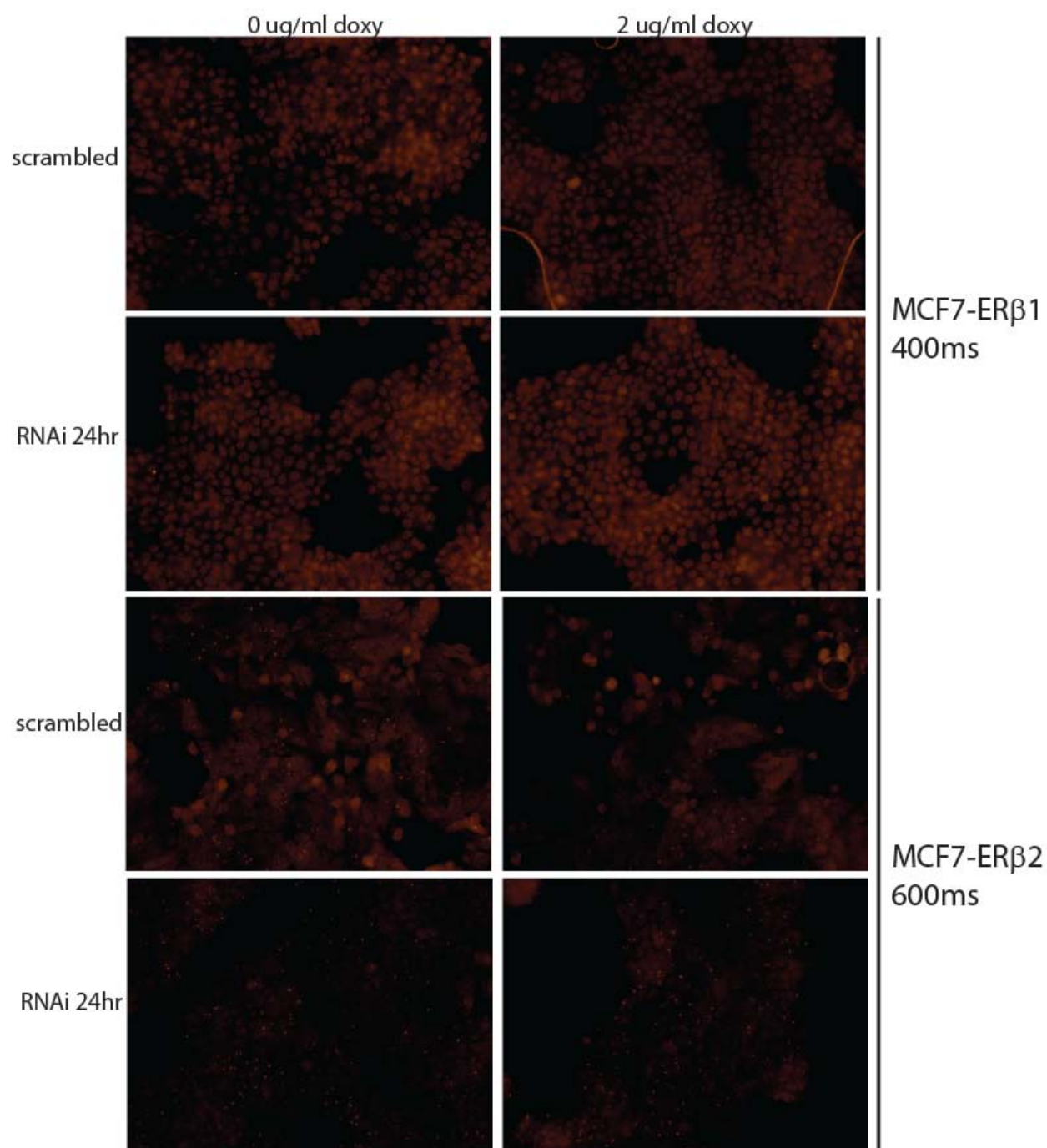
Figure 6. ERβ antibody validation in cell lines. MCF7 cells (endogenous ERβ) were engineered to overexpress ERβ1 or ERβ2 in response to doxycycline. Cells were grown on coverslips and stained using IF with published ERβ1 or ERβ2 specific antibodies (Serotec). Neither antibody appeared to detect an induction of expression **(right panels)** with 2mg/ml doxy. Inhibition of expression with 24hr RNAi treatment (against ERβ total), was not detected with ERβ1 antibody **(2nd row)**, but some decrease in staining was modestly detected with ERβ2 antibody **(4th row)**. Images for each cell line were taken at the same exposure times.
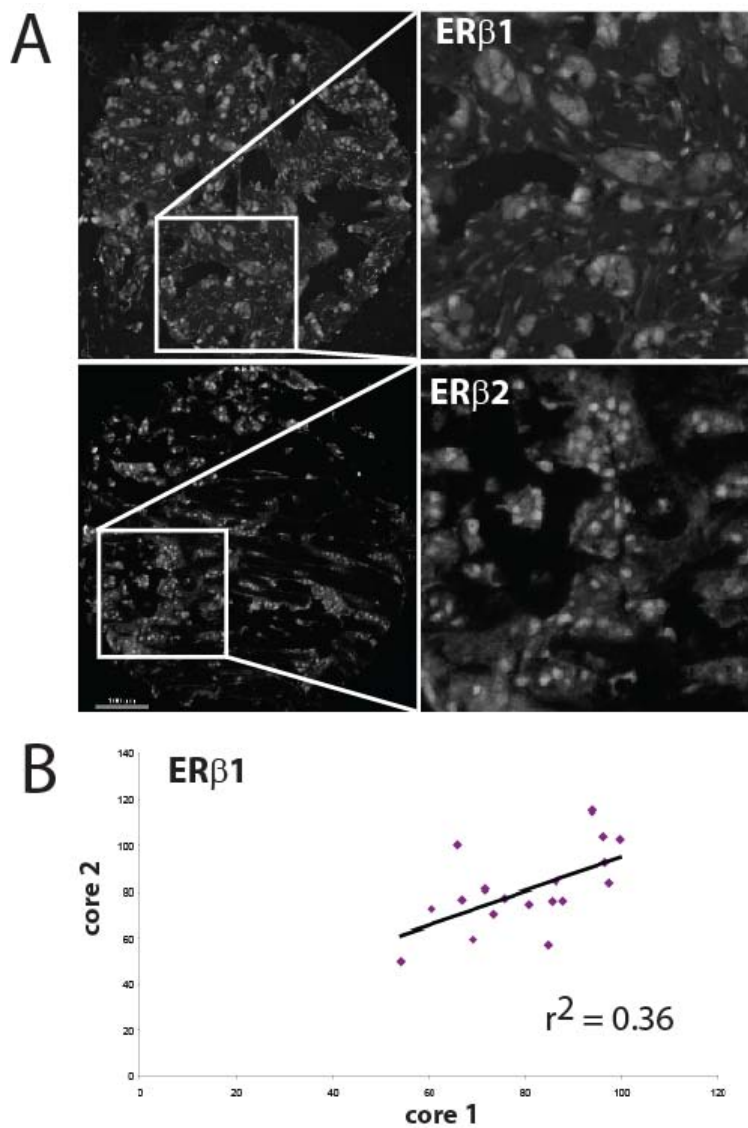
Figure 7. ERβ antibody validation on FFPE clinical breast cancer specimens. Antibodies reported in the literature to detect ERβ isoforms showed non-specific staining when tested by IF on a panel of 40 formalin-fixed, paraffin-embedded (FFPE) breast cancer patients **(A)**. IF expression was quantified with AQUA, and a poor regression was found between duplicate cores from the same patient with ERβ1 antibody (r^2 = 0.36, **B**). The same results for reproducibility were found with ERβ2 (data not shown).
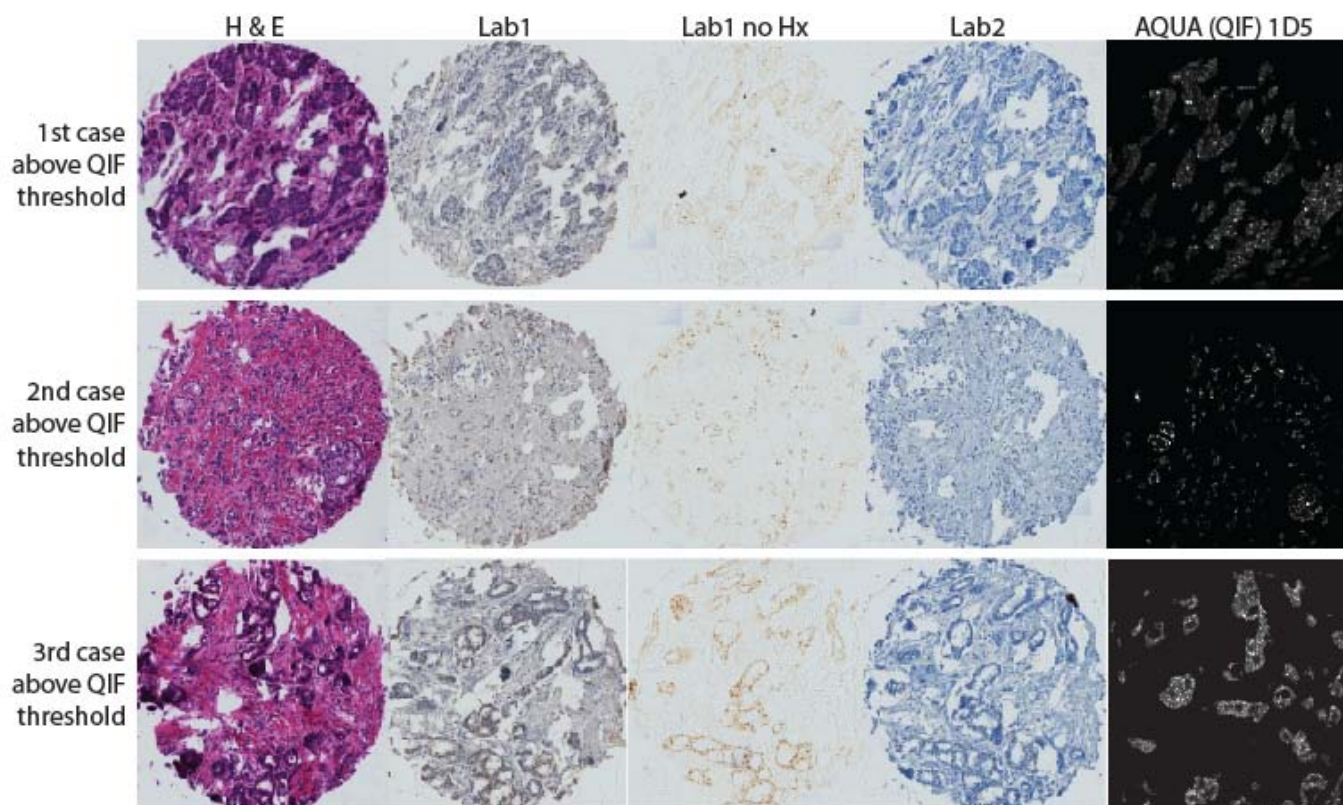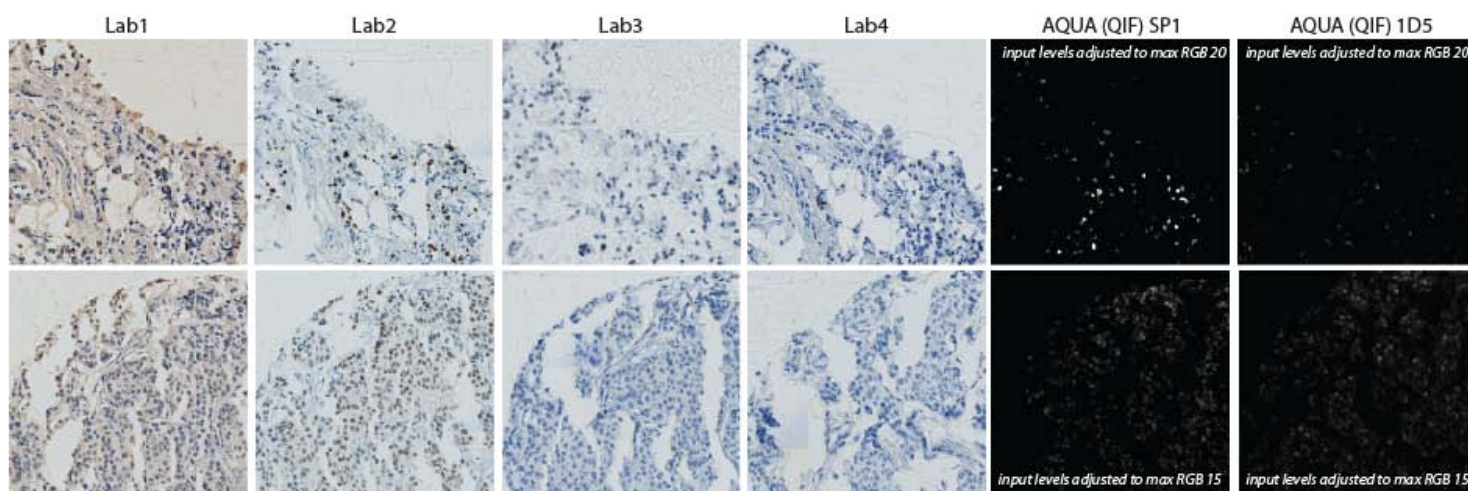
Figure 8. Images of the cases with discordant ER status on the Index TMA. The QIF assay (see last year's progress report, or manuscript attatched in Appendix for detailed description of assay) was performed on the 40 patient controls on the Index TMA and compared to IHC done by routine protocol in two labs as well as in one lab without the hematoxylin (Hx) counterstain. Representative images are shown of the three cases with discordant ER status (2/3 positive by Lab1, 3/3 positive by Lab1 without Hx, 0/3 positive by Lab2, 3/3 positive by QIF) out of the total 31 scoreable spots.

|	AW		DLR		MH		1% average (±stdev)	10% average (±stdev)
	1%	10%	1%	10%	1%	10%		
Lab1 v Lab2	10.3	9.5			18.7	18.7	14.5 (5.9)	14.1 (6.5)
Lab1 v Lab3	11.9	12.5			9.8	9.8	10.9 (1.5)	11.2 (1.9)
Lab1 v Lab4	26.8	27.7			18.7	19.5	22.8 (5.7)	23.6 (5.8)
Lab2 v Lab3	12.4	12.4	17.9	18.3	16.7	16.3	15.7 (2.9)	15.7 (3.0)
Lab2 v Lab4	28.3	28.3	30.5	32.1	28.5	31.8	29.1 (1.2)	30.7 (2.1)
Lab3 v Lab4	17.0	16.8	16.7	18.7	17.7	19.4	17.2 (0.7)	18.3 (1.4)

Table 2. Percent of cases on YTMA 49 with discordant ER status when stained in different labs & scored using different guidelines (10% or 1%). Percentage of cases on Yale TMA 49 cohort (total cases scoreable was 529 by AW, 558 by DLR, 512 by MH) with discordant ER status when comparing routine IHC done by 4 labs (3 clinical, 1 research). The TMAs from each lab were scored for intensity (0-3) and %-positivity by three individuals (DLR, MH certified pathologists, AW graduate student in pathology), and binarized for ER-positivity using the 1% or 10%-positive threshold.
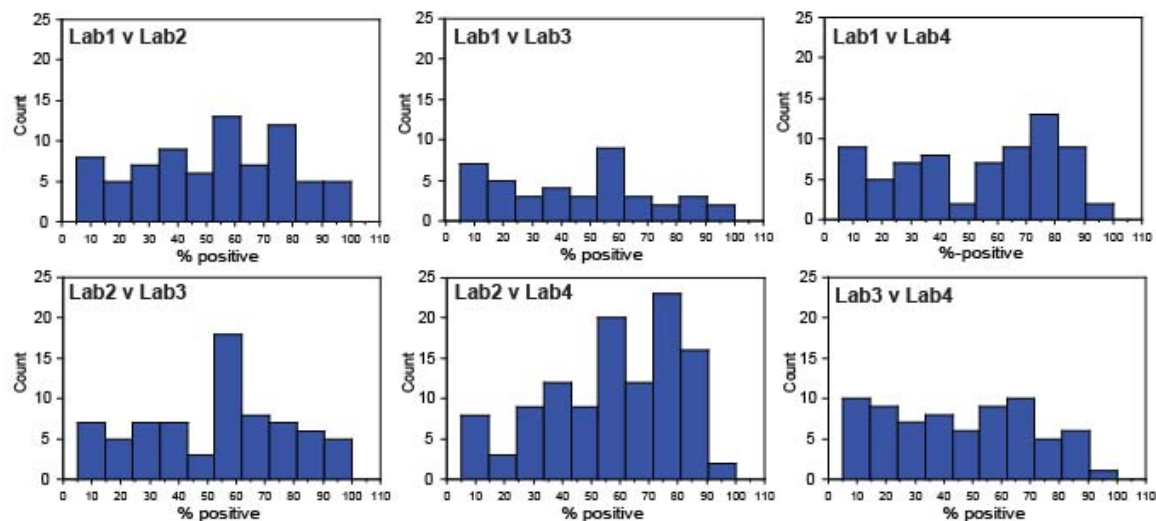
Figure 9. Distribution of %-positive scores in the subset of cases on YTMA 49 with discordant ER status when comparing IHC done by various labs. The distribution of scores for %-positive cells is shown for the subset of cases with discordant ER status lab-to-lab (**see Table 2**). Scores appear evenly distributed across a range from 5% -100%, with the majority well above the 1-10% threshold.

Figure 10. Two examples of the 30% of cases on YTMA 49 with discordant ER status. Representative images of two cases with discordant ER status (**see Table 2**) in YTMA 49. IHC images are shown from each of the four labs as well as AQUA images from QIF performed using SP1 (Thermo/Ventana) or 1D5 (Dako) antibody. QIF images are adjusted to visualize low levels of staining.
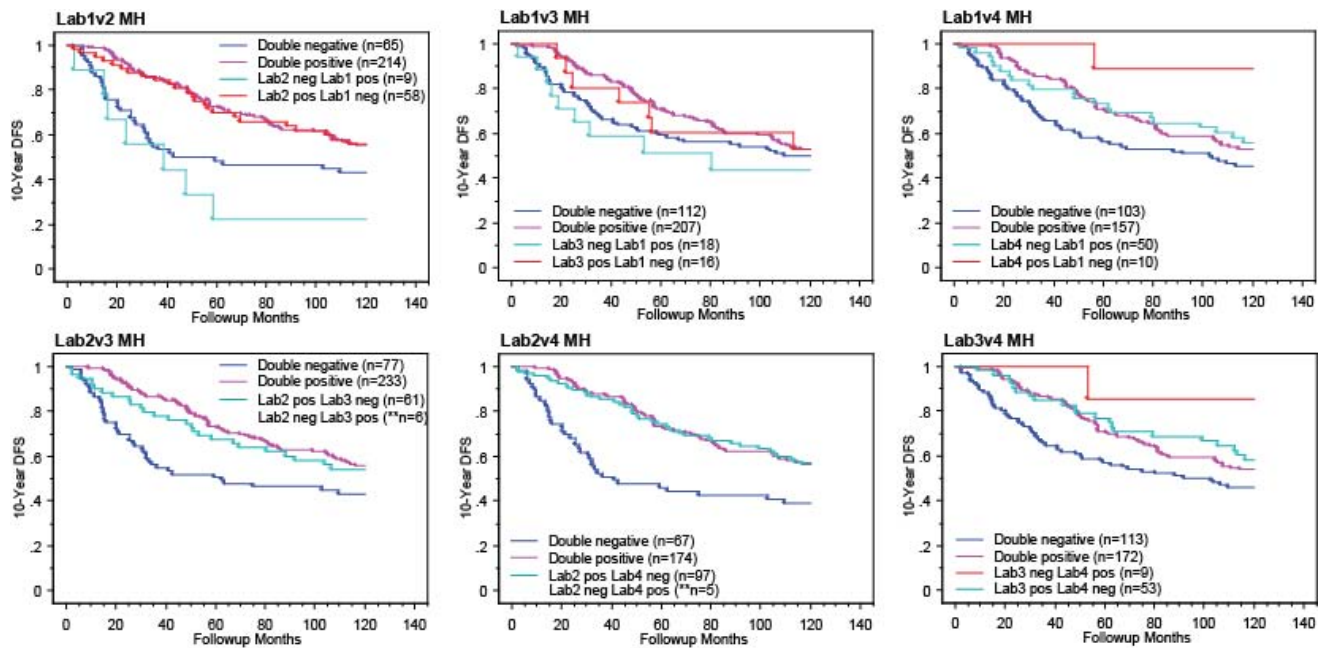
Figure 11. Kaplan-Meier survival analysis of cases on YTMA 49 with concordant & discordant ER status when comparing IHC done in four different labs. Kaplan-Meier disease-specific survival analysis of patients on YTMA 49, stratified by ER status as determined from IHC stain performed in four different labs. Survival curves are only shown for one individual who scored the TMAs (MH), but are similar for all three scorers. Individual curves were eliminated for subgroups with too few patients (n < 9), but the subgroups are still listed.
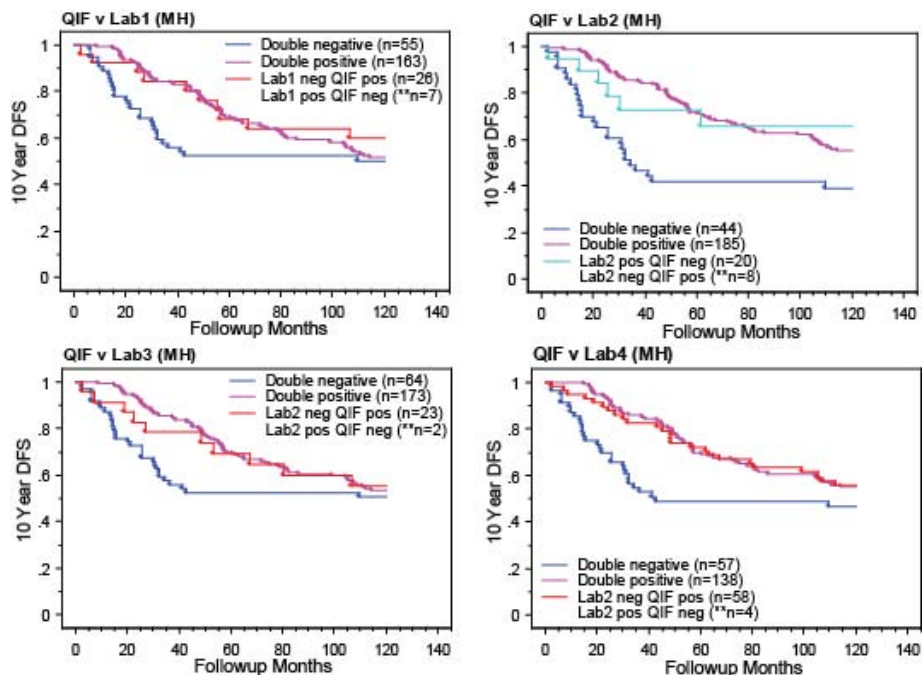
Figure 12. Kaplan-Meier survival analysis of cases on YTMA 49 with concordant & discordant ER status when comparing QIF (AQUA) to IHC done in four different labs. Kaplan-Meier disease-specific survival analysis of patients on YTMA 49, stratified by ER status as determined from IHC stain performed in four different labs compared to QIF analysis with 1D5 antibody (using AQUA and the Index TMA for standardization of ER threshold). Survival curves are only shown for one individual who scored the TMAs (MH), but are similar for all three scorers. Individual curves were eliminated for subgroups with too few patients (n < 9), but the subgroups are still listed.

24
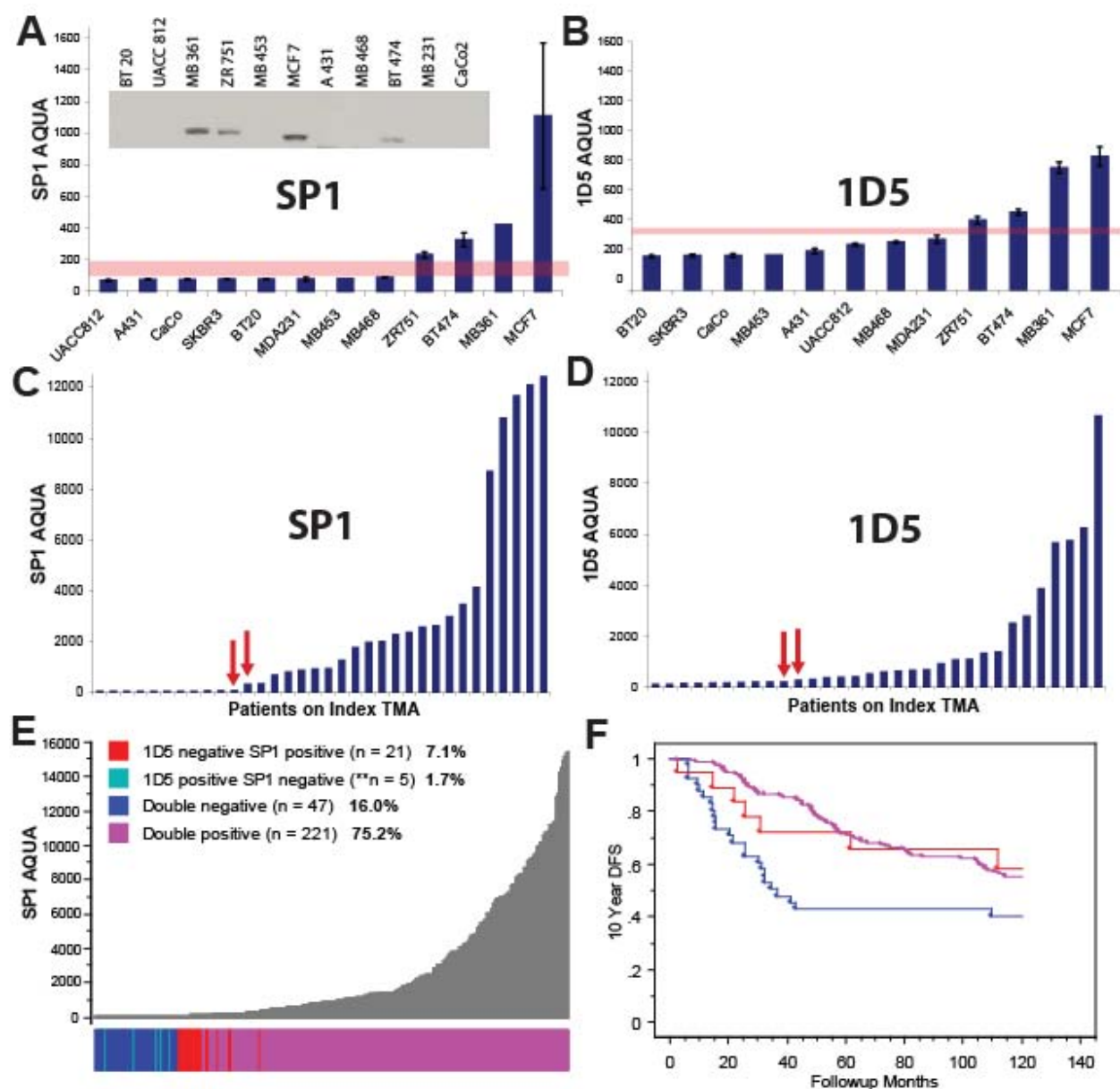
Figure 13. Discordance in ER status on YTMA 49 with SP1 versus 1D5. QIF using AQUA was performed on the Index TMA and YTMA 49 using both SP1 (Thermo/Ventana) and 1D5 (Dako) antibodies. Analysis of cell lines on the Index TMA **(A and B)** and patients on the Index TMA **(C and D)** showed both antibodies had the same threshold for positivity (western blot to confirm positive cell lines shown in **inset, A**), but difference between signal and background (i.e. ER threshold as determined by QIF, red bars in **A and B**, arrows in **C and D**) is greater with SP1 (see also QIF images, **Figure 10**). On YTMA 49, 8.8% of patients had discordant ER status with SP1 vs. 1D5 **(E)**, with the majority (7.1%) of these ER positive with SP1 but ER negative with 1D5. Kaplan-Meier disease-specific survival analysis of the patients on YTMA 49 (stratified as shown in **E**) is shown in **F**. The curve for 1D5 positive / SP1 negative was eliminated in **F** due to small numbers (n < 9).

25

APPENDIX:

Standardization of Estrogen Receptor Measurement in Breast Cancer Suggests False Negative Results are a Function of Threshold Intensity Rather than Percentage of Positive Cells. Welsh AW, Moeder C, Alarid E, Haffty B, Rimm DL. Submitted to JCO, currently under final revisions. (see following pages, 26 in total)

Standardization of Estrogen Receptor Measurement in Breast Cancer Suggests False Negative Results are a Function of Threshold Intensity Rather than Percentage of Positive Cells.

Allison W. Welsh[1], Christopher B. Moeder[1], Sudha Kumar[1], Peter Gershkovich[1], Elaine T. Alarid[2], Malini Harigopal[1], Bruce G. Haffty[3] and David L. Rimm[1]

[1]Department of Pathology, Yale University School of Medicine; [2]McArdle Laboratory for Cancer Research, University of Wisconsin Madison; [3]Department of Radiation Oncology, The Cancer Institute of New Jersey

Correspondence to:

David L. Rimm MD, PhD

Department of Pathology, BML 116

Yale University School of Medicine

310 Cedar Street, PO Box 208023

New Haven, CT 06520-8023

Tel: 203-737-4204

Fax: 203-737-5089

Email: david.rimm@yale.edu

Running Header: Standardization of Estrogen Receptor Measurement
Words: 2948

Abstract:

Purpose

Recent misclassification (false-negative) incidents have raised awareness concerning limitations of immunohistochemistry (IHC) in assessment of Estrogen Receptor (ER) in breast cancer. Here we define a new method for standardization of ER measurement and then examine both the change in percentage and the threshold of intensity (immunoreactivity) to assess sources for test discordance.

Methods

An assay was developed to quantify ER using a control tissue microarray (TMA) and a series of cell lines, where ER immunoreactivity was analyzed by quantitative immunoblotting in parallel with the AQUA method of quantitative immunofluorescence (QIF). The assay was used to assess the ER protein expression threshold in two independent retrospective cohorts from Yale and compared to traditional methods.

Results

Two methods of analysis showed that change in percentage of positive cells, from 10% to 1%, did not significantly affect the overall number of ER+ cases. The standardized assay for ER on two Yale TMA cohorts showed 67.9% and 82.5% of cases above the 2pg/μg immunoreactivity threshold. When compared to pathologist-performed judgment of threshold, we found 9.1% and 19.7% of patients to be QIF+/IHC-, and 4.0% and 0.4% to be QIF-/IHC+, for a total of 13.1% and 20.1% discrepant cases. Assessment of survival for both cohorts showed that QIF-positive, pathologist-negative patients show outcomes more similar to cases with both assays positive.

Conclusion

Assessment of intensity threshold by use of a quantitative, standardized assay on two independent cohorts suggests discordance with current IHC methods in the 10-20% range, where discrepant cases show prognostic outcomes similar to concordant ER-positives.

Introduction

It is widely recognized that the IHC test has significant limitations in accuracy due to a wide range of variables [1]. These issues were highlighted by a recent incident in Canada, which revealed a 40% misclassification rate between local and central laboratories [2] and raised urgent awareness of these existing limitations in ER measurement [3-6]. To address this issue, the American Society of Clinical Oncology and the College of American Pathologists convened an expert panel that ultimately issued a series of guidelines([7]). Most significantly, the guidelines lowered the standard for ER-positivity from 10% to 1%-positive nuclei, but they did not address the issue of intensity or threshold (what actually constitutes a "positive" nucleus). They define positivity as "immunoreactivity… in the presence of expected reactivity of internal (normal epithelial elements) and external controls."

While this may represent the state of the art for assessment of immunoreactivity, it lacks a mechanism for universal standardization. Since amount of ER is scored qualitatively by eye, there is variability and lack of reproducibility between pathologists. Different labs use different antibodies, reagents, and protocols to prepare ER slides for interpretation. To compound the problem, there has been a broad shift to core biopsy over the last few years, so specimens are commonly too small to have "normal epithelial elements" on the same slide. Here we describe a potential method for standardization of ER measurement on a slide. We use quantitative immunofluorescence (QIF), now commercialized as AQUA technology (HistoRx Inc, New Haven, Connecticut) This method calculates marker expression on a continuous scale, using intensity

of pixels, and is shown to be widely applicable for biomarker analysis [8-14].
Previous measurements of ER by AQUA have correlated well with IHC analysis
on tissue from two large clinical trials, as well as predicted response to
Tamoxifen [15] [16].

In an attempt to both quantify and standardize the measurement of ER in
patient tissue, we first sought to define an ER cutpoint with biological and clinical
relevance. This was done using a control TMA (Index array), containing 40
patient controls alongside a panel of cell lines (prepared as tissue and built onto
the TMA). This Index Array is used as a standard and stained alongside every
cohort that is assessed for ER, to allow reproducible selection of the threshold
for positivity. Finally, we used this standardized assay on two independent
archival Yale cohorts, in order to estimate the level of discordance as a function
of intensity threshold (rather than percent-positive) in sample populations.

Methods

All methods are provided in detail in the Supplemental Material.

Cell Line Panel & Culture

A panel of ATCC breast cancer cell lines was chosen to span a range of
ER expression. We also included Puro9 cells (MCF-7 with tetracycline-inducible
ER-alpha overexpression) [17], maintained as six separate cultures (treated with 0,
0.01, 0.1, 0.5, 1, 5 mg/mL doxycycline).

Quantitative Immunoblotting

Amount of ER was quantified (using 1D5 antibody, Dako) as a
concentration (pg ER per µg total protein) for each cell line.

Immunofluorescent staining

TMAs were stained for DAPI, Cytokeratin and ER (1D5 antibody), using a standard protocol developed in our lab. IHC assessment of ER was done by two board-certified pathologists at Yale (MH and DLR,) or at The Cancer Institute of New Jersey, using the 1D5 antibody and standard IHC methods (new 1% cutoff guidelines for YTMA 49, and 10% cutoff for YTMA 130). These IHC assessments were done on the same TMAs used for analysis by the AQUA assay, and thus the same core from each patient.

AQUA Analysis

ER immunofluoresence (IF) was quantified in tumor nuclei using AQUA technology, which was previously developed in the lab.

Patient cohorts

Two large cohorts of archival breast cancer samples from Yale were used: YTMA 49 (diagnosed 1962-1982, n =619) and YTMA 130 (diagnosed 1976-2005, n = 390). Tissues were collected in accordance with consent guidelines in protocol #8219 to Dr. Rimm from the Yale Human Investigation Committee (Institutional Review Board). Clinicopathologic characteristics of both are found in Supplemental Table 1.

Statistical Analysis

All analyses were performed using the StatView software platform. Box plots, ANOVA tests, and Kaplan-Meier survival analyses were performed on each cohort (disease-free survival or recurrence-free survival), and statistical significance assessed using the log-rank test.

Results

Assessment of Discordance as a Function of the Change from 10% to 1% Immunoreactive Cells

Although it has only been a relatively short time since the adoption of the new ASCO/CAP guidelines for percent-positivity at our institution, we have a sufficient volume of patient data to address the effect on ER-positive classification. Using a custom-designed retrospective search of the Yale Copath database, we determined the percentage of total cases called ER-positive by the 10% standard for each year since 2000. We then compared this number to the percentage of cases called positive since April of 2010 (when the 1% standard came into effect). Table 1, using chi square analysis, shows that there is not a significant difference in the percentage of cases called positive using the adopted 1% standard compared to the 10% standard when pairwise comparing cases read in 2010 according to the new standard, to any previous year.

To test this difference in an experimental setting, 3 observers (two pathologists and one student) scored the conventionally-stained TMA according to the new ASCO/CAP guidelines, including both an intensity score and a percentage score. Table 2 shows that there is almost no difference (around 1% of cases) in the percentage of cases called ER-positive using the 10% or 1% cutoff.

Development of an Immunoblot-Standardized Method for Quantification of ER

6

In order to allow reproducible and quantitative selection of an ER cutpoint, we sought to create a control array (which we call the Index TMA), that would serve as a standard curve for ER expression, and include both a panel of cell lines (prepared as patient tissue) as well as 40 patient controls. The goal of using a cell line panel was to perform quantitative western blotting (provides ER measurement as a concentration) in parallel with quantitative IF (provides ER measurement as an AQUA score), in order to create a conversion from AQUA scores to concentrations that could be applied to the 40 patient controls.

For the cell line panel, we chose ATCC breast cancer cell lines representing the range of ER levels. To expand the ER dynamic range so it more closely mirrored that seen in patients, we utilized MCF-7 cells stably transfected with a tetracycline-inducible ER over-expression system (cultured at 0, 0.01, 0.1, 0.5, 1 and 5mg/ml doxycycline) as previously described[17]. ER was measured in this panel of cell lines by quantitative western blot (Figure 1A) alongside a standard curve of recombinant ER (rER), to determine absolute concentration of ER in pg/µg total protein. Cell lines were also prepared as tissue (pelleted, formalin-fixed, paraffin-embedded, and cored) and placed on the Index TMA alongside 40 patient controls, for quantitative IF analysis by AQUA (scores shown in Figure 1B). The same ER antibody (1D5) was used for both western blot and IF analysis. Combining the AQUA and quantitative ER determination from select cell lines, absolute concentrations of ER (in pg/µg) were correlated to ER AQUA scores, and the regression (Figure1C) was used to determine concentrations of ER (pg/µg) from AQUA scores in the cell line panel. Known ER expression in these cell lines allowed us to determine the cutpoint

between the highest ER-negative cell line and the lowest ER-positive cell line to be 2 pg/µg.

This cutpoint was applied to the panel of 40 patient controls on the Index TMA, whose ER concentrations (pg/µg) were calculated from their AQUA scores using the same regression (Figure 1C). There was one patient which did not have sufficient tissue for AQUA analysis, and thus the final panel consisted of 39 patient controls (Figure 1D). We further validated this threshold of 2 pg/µg by eye, contracting the dynamic range of the grayscale (adjusted maximum RGB input level from 255 to 16 using Adobe Photoshop) in order to visualize very low levels of specific nuclear staining, as well as non-specific background. Corresponding images for the highest negative control case (blue arrow in Figure 1D) and the lowest positive control case (pink arrow in Figure 1D) are shown in Figure1E.

This Index TMA is incorporated as a key component of the ER AQUA assay, stained as a control in every experiment, to determine a cutpoint and standardize scores between users, machines and sites. It is assessed for reproducibility with each staining run, and over the course of 8 individual runs has displayed an average coefficient of determination (r^2) of 0.902 (r = 0.950).

Comparison of ER quantification by QIF versus Pathologist Review

In order to determine the effects of a standardized threshold compared to current standard methods, we used our assay to measure ER on two independent retrospective breast cancer cohorts from Yale. For each, ER status was determined as described above (using the Index TMA) and compared to ER status as determined by IHC review (read by two independent pathologists,

where 0 is negative and 1-3 is positive). The first cohort (YTMA 49) is a retrospective collection from Yale consisting of 619 patients, with median follow-up time of 104.1 months (clinicopathological characteristics described in Supplemental Table 1). Due to TMA exhaustion, valid data for ER expression at 2-fold redundancy was obtained on 280 patients.   We saw an overall high concordance between the QIF assay and IHC review (supplemental Figure 1A). Of a total of 252 patients 33 (13.4%) were discordant and 23 (9.1%) were ER-positive by QIF analysis and negative by IHC review (QIF+/IHC-, Table 3).

Quantification of ER revealed a unimodal distribution with 70.7% of cases above the 2 pg/µg threshold and thus defined positive (Figure 2A). The distribution of discordant cases showed that many of them fell around the 2 pg/µg threshold (Figure 2A), as expected. In order to examine the significance of this discordance with respect to patient prognosis, we performed Kaplan-Meier survival analysis using disease-free survival (DFS) as an endpoint. Stratifying patients using both methods of ER analysis (Figure 2B), we found that the patients with discrepant ER status (ER-positive by QIF, negative by IHC) displayed survival behavior that aligned with cases that were ER-positive by both assays (QIF+/IHC+). In order to further validate the 2 pg/µg threshold on this cohort, we visually examined images of ER QIF staining in patients on either side of the cutpoint (Figure 2C). We confirmed specific nuclear staining seen above the threshold at 4.5 pg/µg, in contrast to low levels of non-specific background seen below, at 0 pg/µg.

The second cohort (YTMA 130) is a newer retrospective collection from Yale consisting of 390 patients, 49% of whom had received Tamoxifen, with a

median follow-up time of 80 months (clinicopathological characteristics described in Supplemental Table 1). Of these, 234 patients had valid data on ER status by the QIF assay. Again we saw a strong correlation between IHC review and QIF analysis (Supplemental Figure 1B), but a total of 47 patients (20.1%) still showed discordance, with 98% (46 of 47) of them QIF+/IHC- (Table 3). Representative AQUA/IF images of ER staining for each of these classifications are shown in Supplemental Figure 1C, confirming specific nuclear staining in patients considered positive by QIF analysis but negative by IHC review. Similarly, we saw non-specific background staining in patients who were classified as QIF -/IHC+.

Quantification of ER on this cohort revealed a unimodal distribution with 82.5% of cases above the 2 pg/µg threshold (Figure 3A). Examining the distribution of discordant cases again showed that many were around the threshold, but some were also at the high range of expression. Kaplan-Meier analysis was performed using RFS instead of DFS because data on patient recurrence was available on this cohort, and also because Tamoxifen-treatment reduced the overall number of deaths. Stratification of patients using both methods of ER analysis (Figure 3B), showed that the patients with discordant ER status (QIF+/IHC-) displayed survival behavior that was similar to the double ER-positive population. As we did previously, we visually validated the 2 pg/µg AQUA threshold on patients at either side of the cutpoint (Figure 3C), confirming specific nuclear staining seen at 3.8 pg/µg, but nothing specific detectable at 0.4 pg/µg.

Discussion

The two key findings of this study are 1) the threshold of immunoreactivity appears to be more important that the percentage-positive in generation of discordant or false-negative assays and 2) the standardization method using the QIF assay appears to be more sensitive than the traditional IHC assay, even though the same antibody is used for detection of ER (1D5). In support of the first point, though some pathologists report calling more cases positive as a result of the change in the guidelines, the two data collections examined in this study suggest that false negatives, like those reported in the Canadian incident[2], are unlikely to be due to percentage-positive issues.

False-negative cases may well be a significant problem at other sites around the world as well. Recently-presented data on the ER false-negative rate in the BIG-1-98 population and ALTTO trial also suggested that between 15 and 20% of cases done in local labs may be falsely assigned a negative score. Other studies in the US have much more modest disagreement between centralized versus local laboratories [18] [19], but essentially no labs, in the US or elsewhere, use a standard curve to assess the ER detection threshold. The current standard in most labs is to use a single strongly positive example case as a control for stainer runs. Other labs rely on intrinsic controls provided by adjacent normal ducts. Neither of these methods specifically assesses the threshold of positivity.

The second key finding of this study is that the use of a standardized method results in a reproducible system for assessment of that threshold. Furthermore, it reveals a threshold that by QIF appears to be more sensitive than traditional IHC. This may be due to the use of the hematoxylin counterstain that, when applied too heavily, can obscure faint staining, as has been previously described for other tumors [20]. Examples of 2 discordant cases that were QIF+/IHC- are shown in Supplemental Figure 2. Some automated technologies claim to be able to "unmix" the colors, and they may have similar capacity and sensitivity. However, to our knowledge, a head to head comparison has not yet been done.

There are a number of limitations in the conclusions that can be drawn from this study. Perhaps the most important is that we are unable to determine ground truth for estrogen receptor status. Although we can assess test discordance, and compare discordant cases to concordant cases with respect to survival, we have no absolute way of determining the true ER expression status of each patient. The best method to adjudicate this would be response to endocrine therapy. That information is not available for this study, although studies are planned to test this assay in clinical trial specimens where that information is available.

The assay we developed represents our best attempt to accurately measure ER protein in tissue, but any assay can only measure protein that is present on the slide. Pre-analytic factors, most significantly cold ischemic time, can decrease the amount of ER epitope present on the slide, and account for some level of misclassification in the clinical setting [1]. However, in this study,

both assays were performed on the same tissue specimens, thus pre-analytic variation is unlikely to contribute to the observed discordance. Another limitation of this study is that the cohort analyses were done on TMAs rather than whole sections as used in the clinical setting. While TMAs have been shown to be representative, they may represent a limitation with respect to assessment of sufficient area. TMAs may also represent a limitation in that the heterogeneity seen in a tissue section is unlikely to be completely represented in a TMA. In cases of discordance distant from the threshold, the cause could be tumor heterogeneity.

In this study, our goal was to derive a biologically-relevant cutpoint and a method of standardization that could be used in clinical labs. Using cell lines allowed us to convert patient ER expression to an absolute concentration within a field of view. An absolute concentration, along with a confidence interval for measurement, is a standard readout for many laboratory tests based on fluid specimens, and thus a reasonable goal for ER. The use of cell lines may be a good future universal standard. However, we have found that, even if authenticated, they can show variable expression as a function of confluence, passage number, and other variables yet-to-be determined. Studies are underway in the lab to develop alternative universal standards. Although not perfect, we believe the best current standard can be derived from a set of index patients in conjunction with a standardized set of cell lines. The Index TMA in this paper included 39 patients spanning the range of ER expression, represented cases around the threshold, and showed strong run-to-run reproducibility (r > 0.9). It is a good example of a standard array that could be

processed with each stainer run to assure reproducibility around the ER threshold.

Overall, our results suggest that use of a standardized, quantitative, IF-based assay has the ability to significantly improve the way ER status is evaluated, overcoming the limitations of IHC by providing a method for reproducible assessment of the threshold. Furthermore, they suggest potential biological relevance for low levels of ER expression, and reinforce our need to adopt a standardized assay that can discern this subtle, but potentially-important phenomenon. The AQUA method for analysis of patients specimens has now be implemented by a CLIA lab in efforts to offer a more accurate and reproducible test for ER, PR and Her2.

Figure Legends

Figure 1. Method for quantification of ER using an immunoblot-standardized AQUA assay

ER was measured in a panel of cell line controls by western blot (A, 1D5 antibody, Dako) alongside a standard curve of recombinant ER (rER) to determine absolute concentration in pg/µg total protein. Cell lines included Puro9 cells, which are MCF-7s with doxycycline-induced overexpression of ER (0, 0.01, 0.1, 0.5, 1 and 5 mg/ml doxy). Cell lines were also pelleted, cored and placed on the Index TMA for IF & AQUA analysis. Absolute concentrations of ER (pg/µg) were correlated to ER expression by IF (AQUA score using 1D5, B), and the regression (C) was used to convert AQUA scores to concentrations of ER (pg/µg) in the set of patient controls present on the same Index TMA (pg/µg distribution shown in D). Immunofluorescent AQUA images of ER in the highest negative control case (blue arrow in D) and the lowest positive control case (pink arrow in D) are shown in E to validate the cutpoint. Cytokeratin (CK) was used as a mask to define regions of tumor. For ER, we contracted the dynamic range of the grayscale (adjusted maximum RGB input level from 255 to 16 using Adobe Photoshop) in order to visualize very low levels of specific nuclear staining as well as non-specific background. (ER, Estrogen Receptor; AQUA, Automated Quantitative Analysis)

Figure 2. Discordant classification of ER status in YTMA 49 cohort

A) ER status was determined by IF & AQUA analysis in a Yale retrospective breast cancer cohort YTMA 49 (diagnosed 1953-1983, clinicopathological

characteristics in Supplementary Table 1) and compared to ER status as determined by IHC, read by two certified pathologist (MH and DLR) using the current 1%-positive nuclei cutoff guidelines.   A distribution of ER by AQUA (pg/µg standardized as shown in Figure 1) is shown where each case is color-coded in the bar below, according to its ER status by both AQUA and IHC. B) Kaplan-Meier curves show 10-year DFS, where patients are grouped according to the classifications shown in A.   The AQUA-/IHC+ group (n=10) was excluded from survival analysis on account of small size and insufficient power. C) To confirm and further validate the AQUA cutpoint of 2 pg/µg on this cohort, representative IF images of ER staining for patients on either side of the cutpoint are shown (right panels). Cytokeratin (CK) was used as a mask to define regions of tumor (green, left panels). For ER, we contracted the dynamic range of the grayscale (adjusted maximum RGB input level from 255 to 16 using Adobe Photoshop) in order to visualize very low levels of specific nuclear staining as well as non-specific background. (ER, estrogen receptor; IF, immunofluorescence; AQUA, Automated Quantitative Analysis; IHC, immunohistochemistry; DFS, disease-free survival).

Figure 3. Discordant classification of ER status in YTMA 130 cohort

A) ER status was determined by IF & AQUA analysis in a second Yale retrospective breast cancer cohort YTMA 130 (diagnosed between 1976-2005, clinicopathological characteristics in Supplementary Table 1) and compared to ER status as determined by IHC using the 10%-positive nuclei cutoff guidelines. A distribution of ER by AQUA (pg/µg standardized as shown in Figure 1) is

shown, where each case is color-coded according to its ER status by both AQUA

and IHC.   B) Kaplan-Meier curves show 10-year RFS, where patients are

grouped according to the classifications shown in A. The AQUA-/IHC+ group

(n=1) was excluded from survival analysis on account of its size and insufficient

power. C) The AQUA cutpoint of 2 pg/µg was further validated on this cohort by

examining representative immunofluorescent images of ER staining for patients

on either side of the cutpoint (right panels). Cytokeratin (CK) was used as a

mask to define regions of tumor (green, left panels). For ER, we contracted the

dynamic range of the grayscale (adjusted maximum RGB input level from 255 to

16 using Adobe Photoshop) in order to visualize very low levels of specific

nuclear staining as well as non-specific background. (ER, estrogen receptor; IF,

immunofluorescence; AQUA, Automated Quantitative Analysis; IHC,

immunohistochemistry; RFS, recurrence-free survival).

Table 1. Number of invasive breast carcinoma cases diagnosed as ER-positive at Yale-New Haven Hospital from 2000-2010.

Year	Total number of invasive carcinoma cases with ER results	Total number of invasive carcinoma cases with ER-positive results	Percent ER-positive	Chi-square p values for pairwise comparison with 2010 data
2000	246	189	76.83%	0.29
2001	268	212	79.10%	0.60
2002	264	196	74.24%	0.09
2003	298	226	75.84%	0.18
2004	332	266	80.12%	0.79
2005	455	342	75.16%	0.11
2006	491	406	82.69%	0.64
2007	497	395	79.48%	0.64
2008	502	411	81.87%	0.82
2009	550	450	81.82%	0.83
From April 2010	180	146	81.11%	-

ER = Estrogen Receptor. Data from 2010 includes only April through August 31. Note that over the last 10 years there has been a statistically significant trend toward increase in ER in the population seen at Yale (Mantel-Haenszel Chi-Square p=0.0036).

Table 2. Number of invasive breast carcinoma cases scored as ER-positive on Yale TMA 49.

Scorer	Total number invasive carcinoma cases with ER results	Total number cases scored ER-positive using 10% cutoff	Percent cases scored ER-positive using 10% cutoff	Total number cases scored ER-positive using 1% cutoff	Percent cases scored ER-positive using 1% cutoff
DLR	526	312	59.31%	318	60.46%
MH	462	293	63.42%	293	63.42%
AWW	502	335	66.73%	340	67.73%

ER = Estrogen Receptor. Excluded cases were unscoreable due to insufficient tumor, infiltration, or out-of-focus tissue. DLR and MH are board-certified pathologists; AWW is a graduate student in Pathology.

Table 3. Comparison of ER status by IHC review versus AQUA assay for YTMA 49 and YTMA 130

ER status by AQUA (positive > 2 pg/µg, negative < 2 pg/µg)	ER status by IHC review (positive = 1-3, negative = 0)	YTMA 49 (1962-1982) N (%)	YTMA 130 (1976-2005) N (%)
positive	positive	148 (58.7)	147 (62.8)
positive	negative	23 (9.1)	46 (19.7)
negative	positive	10 (4.0)	1 (0.4)
negative	negative	71 (28.2)	40(17.1)
	Total	**252**	**234**

ER = Estrogen Receptor, IHC = immunohistochemistry, AQUA = Automated Quantitative Analysis.

References:

1.   Gown AM: Current issues in ER and HER2 testing by IHC in breast cancer. Mod Pathol 21 Suppl 2:S8-S15, 2008

2.   Hede K: Breast cancer testing scandal shines spotlight on black box of clinical laboratory testing. J Natl Cancer Inst 100:836-7, 844, 2008

3.   Allred DC: Commentary: hormone receptor testing in breast cancer: a distress signal from Canada. Oncologist 13:1134-6, 2008

4.   Matthews A: Bad cancer tests draw scrutiny, Wall Street Journal, 2008, pp 1-2

5.   Allison KH: Estrogen receptor expression in breast cancer: we cannot ignore the shades of gray. Am J Clin Pathol 130:853-4, 2008

6.   Harris L, Fritsche H, Mennel R, et al: American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. J Clin Oncol 25:5287-312, 2007

7.   Hammond ME, Hayes DF, Dowsett M, et al: American Society of Clinical Oncology/College Of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. J Clin Oncol 28:2784-95, 2010

8. McCabe A, Dolled-Filhart M, Camp RL, et al: Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. J Natl Cancer Inst 97:1808-15, 2005

9. Berger AJ, Davis DW, Tellez C, et al: Automated quantitative analysis of activator protein-2alpha subcellular expression in melanoma tissue microarrays correlates with survival prediction. Cancer Res 65:11185-92, 2005

10. Camp RL, Dolled-Filhart M, King BL, et al: Quantitative analysis of breast cancer tissue microarrays shows that both high and normal levels of HER2 expression are associated with poor outcome. Cancer Res 63:1445-8, 2003

11. Dolled-Filhart M, McCabe A, Giltnane J, et al: Quantitative in situ analysis of beta-catenin expression in breast cancer shows decreased expression is associated with poor outcome. Cancer Res 66:5487-94, 2006

12. Giltnane JM, Ryden L, Cregger M, et al: Quantitative measurement of epidermal growth factor receptor is a negative predictive factor for tamoxifen response in hormone receptor positive premenopausal breast cancer. J Clin Oncol 25:3007-14, 2007

13. Gustavson MD, Bourke-Martin B, Reilly D, et al: Standardization of HER2 immunohistochemistry in breast cancer by automated quantitative analysis. Arch Pathol Lab Med 133:1413-9, 2009

14. Moeder CB, Giltnane JM, Harigopal M, et al: Quantitative justification of the change from 10% to 30% for human epidermal growth factor receptor 2 scoring in the American Society of Clinical Oncology/College of American Pathologists guidelines:

tumor heterogeneity in breast cancer and its implications for tissue microarray based assessment of outcome. J Clin Oncol 25:5418-25, 2007

15.   Bartlett JMS BC, Robson T et al: The TEAM Trial Pathology Study Identifies Potential Prognostic and Predictive Biomarker Models for Postmenopausal Patients Treated with Endocrine Therapy. 32nd Annual San Antonio Breast Cancer Symposium Abstract 75, 2009

16.   Harigopal M, Barlow WE, Tedeschi G, et al: Multiplexed assessment of the Southwest Oncology Group-directed Intergroup Breast Cancer Trial S9313 by AQUA shows that both high and low levels of HER2 are associated with poor outcome. Am J Pathol 176:1639-47, 2010

17.   Fowler AM, Solodin N, Preisler-Mashek MT, et al: Increases in estrogen receptor-alpha concentration in breast cancer cells promote serine 118/104/106-independent AF-1 transactivation and growth in the absence of estrogen. Faseb J 18:81-93, 2004

18.   Ma H, Wang Y, Sullivan-Halley J, et al: Breast cancer receptor status: do results from a centralized pathology laboratory agree with SEER registry reports? Cancer Epidemiol Biomarkers Prev 18:2214-20, 2009

19.   Collins LC, Marotti JD, Baer HJ, et al: Comparison of estrogen receptor results from pathology reports with results from central laboratory testing. J Natl Cancer Inst 100:218-21, 2008
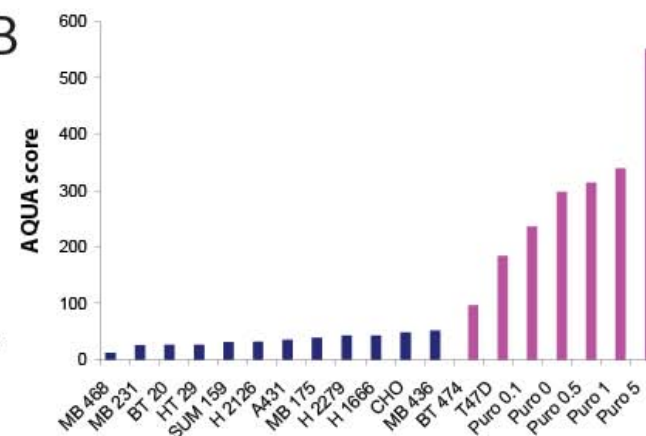
20.   Rothberg BE, Moeder CB, Kluger H, et al: Nuclear to non-nuclear Pmel17/gp100 expression (HMB45 staining) as a discriminator between benign and malignant melanocytic lesions. Mod Pathol 21:1121-9, 2008
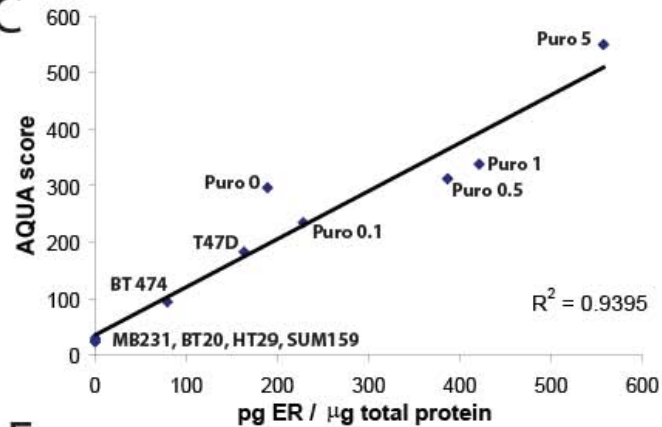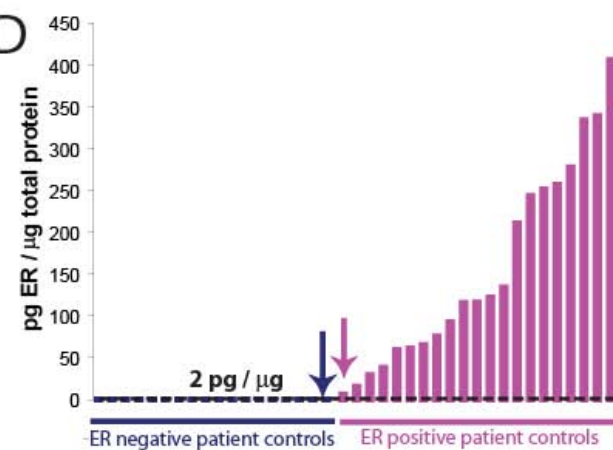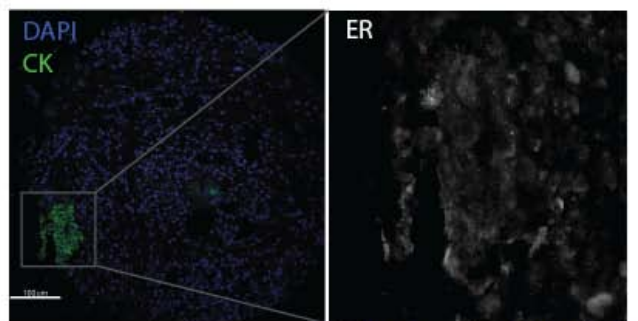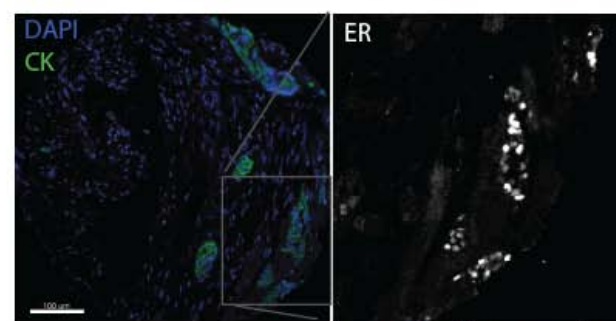
Figure 1

Figure 2

Figure 3